

Cristina Rosalinda Mendes Ventura

**Os Modelos de Exposição Necessários à Aquisição de Publicidade no
Sector Televisivo**

**Universidade Fernando Pessoa
Porto 2015**

Cristina Rosalinda Mendes Ventura

**Os Modelos de Exposição Necessários à Aquisição de Publicidade no
Sector Televisivo**

**Universidade Fernando Pessoa
Porto 2015**

@2015
Cristina Rosalinda Mendes Ventura
“TODOS OS DIREITOS RESERVADOS”

Cristina Rosalinda Mendes Ventura

**Os Modelos de Exposição Necessários à Aquisição de Publicidade no
Sector Televisivo**

Tese apresentada à Universidade Fernando Pessoa
como parte dos requisitos para obtenção do grau de doutor em
Ciências da Informação, ramo Jornalismo e Estudos Mediáticos
sob a orientação do Prof. Doutor Galvão dos Santos Meirinhos

RESUMO

CRISTINA ROSALINDA MENDES VENTURA: Os Modelos de Exposição Necessários à Aquisição de Publicidade no Sector Televisivo (Sob orientação do Prof. Doutor Galvão dos Santos Meirinhos)

O investimento publicitário no sector televisivo depende do desenvolvimento de modelos de “*ratings*” ou da identificação de abordagens metodológicas alternativas de previsão da exposição televisiva. Avaliámos o contributo da Análise Simbólica e do *Data Mining* para a construção dos modelos quantitativos de exposição, que servem de suporte à actividade de planeamento de *media*.

Nas bases de audimetria consta informação com uma considerável capacidade explicativa da evolução dos *ratings* que pode alcançar os 90%. Porém, o potencial predictivo das análises univariadas e multivariadas de Regressão linear e não linear é consideravelmente menor, situando-se no máximo no intervalo 70%-80%. Foram testadas determinadas metodologias de Redes Neurais (MLP e RBF), Árvores de Regressão (CART e CHAID), IBL, segmentação e *clustering* das séries temporais e modelos locais de Regressão.

A construção de modelos explicativos dos comportamentos “estruturais” de consumo televisivo, permitiu verificar que no painel existe uma reduzida a moderada duplicação das audiências e que a totalidade dos comportamentos de lealdade está presente, existindo alguma tendência para a especialização das audiências.

O desenvolvimento de um modelo explicativo estrutural da exposição televisiva demonstra os múltiplos contextos de exposição intencional e não intencional e fundamenta uma proposta alternativa de construção dos modelos de exposição, recorrendo a metodologias simbólicas, ao *Data Mining* Sequencial, Temporal, Multirrelacional e a algoritmos Bayesianos e de Regressão não linear, que é aplicável nos contextos de maior irregularidade dos dados de *ratings* ou quando novos conteúdos são transmitidos.

Para os segmentos que apresentam uma exposição fortemente irregular é proposta a construção de Regras de associação e sequenciais que vão permitir a identificação dos suportes mais adequados à divulgação da mensagem publicitária, com a posterior construção de Redes Bayesianas e de Regras de Classificação multirrelacionais para reduzir a incerteza dos resultados em determinado período. Quando existem hábitos de consumo televisivo poderá ser suficiente recorrer ao *Data Mining* Sequencial, a modelos Binomiais Logísticos ou à Classificação de Bayes.

No contexto de transmissão de eventos desportivos devemos recorrer às Regras Temporais que permitem identificar informação relevante nas séries temporais multivariadas de “*ratings*”, viabilizando uma melhor negociação com as estações televisivas.

Palavras-chave: Audiências, Televisão, *Data Mining*, Previsão, Análise Simbólica, Planeamento de *Media*.

ABSTRACT

CRISTINA ROSALINDA MENDES VENTURA: TV Exposure Models for Media Planning
(Under the orientation of Prof. Dr Galvao dos Santos Meirinhos)

Television advertising investment depends on the development of *ratings* models or on the identification of alternative methodological approaches for the prediction of television exposure. In this research study, we evaluate the contribution of Symbolic Analysis and *Data Mining* for the construction of quantitative exposure models, which support the activity of *media* planning.

According to the results attained, *ratings* databases contain information with a considerable explanatory capacity on the evolution of commercial *ratings*, which can reach up to 90%. However, the predictive potential of univariate and multivariate Linear Regression models and non-linear analysis is considerably lower and in general drops in the 70% -80% range. Certain methodologies were tested within the Neuronal Networks field (MLP and RBF), Regression Trees (CART and CHAID), IBL, segmentation and *clustering* of time series and Local Regression models.

The construction of explanatory models for television “structural” consumption behaviours allowed us assessing that the panel presents reduced audience duplication *ratings* but all of the loyalty behaviours are present and there is a trend towards the specialisation of TV audiences.

The development of a structural explanatory television exposure model demonstrates the multiple contexts of intentional and unintentional TV exposure and justifies an alternative proposal for the construction of exposure models, using symbolic methodologies, Temporal, Sequential and Multi-relational *Data Mining* and Bayesian algorithms and Non-Linear

Regression, which is most suited in the contexts of a higher irregularity of *Ratings* data or when new content is broadcasted.

For audience segments which exhibit stronger irregular patterns, the construction of association or sequence rules is proposed. These rules will allow the identification of the most appropriate commercial spots for the broadcasting of the advertising message, with the subsequent construction of Bayesian Networks and Multi-Relational Regression Rules so as to reduce the uncertainty of the results over a given period. When viewers have television consumption habits, it may be sufficient to use Binomial Logistic models and Data Mining Sequential models or Bayes classification.

In the context of the broadcast of sports events, there is a great difficulty in the construction of causal models. Therefore, we must turn to Temporal Rules in order to identify relevant information in the multivariate *ratings* time series, enabling a better negotiation with the TV stations.

Keywords: Audiences, Television, *Data Mining*, Forecasting, Symbolic Analysis, *Media Planning*.

RÉSUMÉ

CRISTINA ROSALINDA MENDES VENTURA: Les modèles d'exposition nécessaires à l'acquisition de Publicité au Secteur de la Télévision (Sous la supervision du Prof. Dr Galvão dos Santos Meirinhos)

L'investissement au publicité au secteur de la télévision dépend du développement des modèles de "ratings" ou de l'identification de plusieurs approches alternatives de prévision de l'exposition à la télévision. On a évalué le contribut de l'Analyse Symbolique et du Data Mining à fin de créer des modèles quantitatifs d'exposition qui supportent l'activité de planification du media.

Aux bases de l'audiométrie on trouve l'information avec une capacité explicatif considérable sur l'évolution des ratings qui peut atteindre un pourcentage de 90%. Cependant, le potentiel de pronostiquer les analyses univariées et multivariées da la Régression Linéaire et non Linéaire est considerablement inférieur et se situe dans un intervalle 70%-80% maximum. On a examiné certaines méthodologies des réseaux de neurones (MLP et RBF), arbres de régression (CART e CHAID), IBL, segmentation et clustering des séries chronologiques et des modèles locales de Régression.

La création des modèles explicatifs des comportements "structurals" de consommation de télévision a montré qui au panneau existe une duplication des audiences faible à modérée et que tous les comportements de loyauté sont présents et qu'il ya une certaine tendance pour la spécialization des audiences.

Le développement d'un modèle explicatif structural de l'exposition à la télévision montre les contextes variés de l'exposition intentionnel et non intentionnel et soutient une suggestion alternative de création des modèles de exposition, donnant la possibilité de utilisation des méthodologies symboliques, le Data Mining Séquentiel, Temporel, Multirrelacional et algorithmes bayésiens et de Régression non linéaire, qui sont appliqués dans les contextes plus irrégulières des ratings ou quand les nouveaux contenus sont transmis.

Pour les segments qui présentent une exposition beaucoup irrégulière on propose la création des règles de association et séquentielles qui permettront l'identification des supports plus convenables à la divulgation du message publicitaire, avec la création en arrière des règles bayésiens et des règles de classification multirrelationals à fin de réduire l'incertitude des résultats dans un période déterminé. Quand on existe les habitudes de consommation de la télévision sera suffisant utiliser le Data Mining Séquentiel, les modèles Logistiques Binominales ou la classification de Bayes

Au contexte de transmission des événements sportifs on doit appliquer les Règles Temporelles qui identifient l'information plus important dans les séries chronologiques multivariées des "ratings", et qui permet une meilleure négociation avec les chaînes de télévision.

Mots-Clés: Audiences, Télévision, Data Mining, Prévion, Analyse Symbolique, Planification de media

AGRADECIMENTOS

Gostaria de agradecer a todos os que estiveram presente no decorrer deste período, e que de alguma forma contribuíram para a realização da presente tese.

Ao Prof. Dr. Galvão Meirinhos, Prof. Dr. Fernanda Nogueira e ao Prof. Dr. Albino Lopes agradeço o seu apoio para viabilizar a conclusão deste trabalho.

Aos meus pais, avós, amigos e ao Prof. Dr. Amador Cernuda o meu especial obrigado pelo seu contínuo apoio e incentivo no decorrer destes anos.

A presente tese procurou contribuir para a actividade de planeamento de *media*, consequentemente a sua realização não seria viável sem recorrer aos dados de audimetria. Agradeço à agência de planificação de meios, Tempo-Omd, me ter permitido recolher nas suas instalações alguns dados de audimetria.

A temática em análise e os resultados que podem ser alcançados, certamente, que dependem fortemente dos recursos que se encontram acessíveis, nomeadamente os informáticos. Agradeço, a este nível, a colaboração do Instituto Nacional de Estatística de Lisboa e dos diversos autores internacionais que disponibilizaram alguns dos algoritmos de Análise Simbólica que foram utilizados na realização das análises.

ÍNDICE

CAPÍTULO I. INTRODUÇÃO	1
I.1. A PROBLEMÁTICA DA INVESTIGAÇÃO: OS MODELOS DE EXPOSIÇÃO NA ACTIVIDADE DE PLANEAMENTO DE <i>MEDIA</i>	2
I.2. A RELEVÂNCIA DO TEMA DE INVESTIGAÇÃO.....	9
I.3. A QUESTÃO DE PARTIDA.....	17
I.4. AS PRINCIPAIS CONTRIBUIÇÕES DA TESE.....	22
I.5. A ESTRUTURA DA TESE.....	32
 CAPÍTULO II. OBJECTIVOS E QUESTÕES A DESENVOLVER.....	45
II.1. MOTIVAÇÃO E OBJECTIVO GERAL.....	45
II.2. OBJECTIVOS ESPECÍFICOS.....	48
II.3. AS QUESTÕES OPERACIONAIS EM ANÁLISE.....	57
 CAPÍTULO III. FUNDAMENTAÇÃO TEÓRICA.....	69
III.1. INTRODUÇÃO À ACTIVIDADE DE PLANEAMENTO DE <i>MEDIA</i>	69
III.1.1. As etapas que integram a construção de uma campanha publicitária.....	70
III.1.2. A implementação do planeamento de <i>Media</i>	71
III.1.3. Os modelos de <i>Media</i> : Os modelos de exposição, optimização e de avaliação dos efeitos da publicidade.....	73
III.1.4. O plano de <i>Media</i>	77
III.1.4.1. Os indicadores de desempenho do plano.....	78

III.1.4.2. Métodos de construção dos planos e sua avaliação.....	82
III.2. OS MODELOS DE EXPOSIÇÃO NO PLANEAMENTO DE <i>MEDIA</i>	83
III.2.1. Os modelos que antecederam o desenvolvimento da audimetria.....	83
III.2.2. O contributo da audimetria para construção dos modelos de exposição.....	84
III.2.3. As metodologias utilizadas na construção dos modelos de exposição: Vantagens e limitações.....	86
III.3. CONSIDERAÇÕES SOBRE AS ANÁLISES DE PREVISÃO NAS CIÊNCIAS SOCIAIS: AS CARACTERÍSTICAS DOS MODELOS DE EXPOSIÇÃO.	94
III.4. O CONTRIBUTO DOS MODELOS EXPLICATIVOS DO CONSUMO TELEVISIVO E DOS <i>RATINGS</i> PARA A CONSTRUÇÃO DOS MODELOS DE EXPOSIÇÃO.....	103
III.4.1. Os modelos explicativos da exposição televisiva.....	105
III.4.2. A análise dos comportamentos estruturais.....	109
III.4.3. O conceito de hábito de consumo televisivo.....	111
III.4.4. Os modelos de <i>Ratings</i> e as perspectivas explicativas da exposição.....	112
III.5. A INVESTIGAÇÃO REALIZADA NO SECTOR TELEVISIVO E ENQUADRAMENTO DA TEMÁTICA PROPOSTA.....	114
III.5.1. As temáticas em análise no sector TV.....	114
III.5.2. A investigação académica e empresarial.....	115
III.5.3. As metodologias aplicadas aos estudos de audiências.....	116
III.5.4. Estado da arte dos modelos de “ <i>Ratings</i> ”.....	119
III.6. ESTADO DA ARTE DOS MODELOS QUANTITATIVOS DE EXPOSIÇÃO TELEVISIVA.....	122
III.6.1. Os modelos de exposição ao nível dos indicadores agregados ou “Gross measures” de audimetria.....	124
III.6.1.1. Os modelos de exposição univariados.....	125

III.6.1.2. Os modelos de exposição multivariados programáticos.....	126
III.6.1.3. Os modelos de exposição comercial.....	129
III.6.2. Os modelos de exposição ao nível dos indicadores acumulados de audimetria: A duplicação das audiências.....	133
III.6.2.1. Os efeitos de herança entre suportes.....	134
III.6.2.2. A Lealdade programática.....	137
III.6.2.3. A lealdade ao canal.....	138
III.6.2.4. As exposições repetidas.....	140
III.6.2.5. A avaliação da interacção entre as diversas vertentes de análise da duplicação das audiências.....	140
III.6.2.6. Os modelos de exposição por previsão da duplicação das audiências...	141
III.6.3. As abordagens alternativas de construção dos modelos de exposição.....	141
 CAPÍTULO IV. METODOLOGIA.....	 145
IV.1. INTRODUÇÃO.....	145
IV.2. MODELO DE INVESTIGAÇÃO.....	148
IV.3. DESCRIÇÃO DA TECNOLOGIA DE AUDIMETRIA.....	150
IV.4. DESCRIÇÃO DAS BASES DE DADOS.....	152
IV.5. HORIZONTE TEMPORAL.....	153
IV.6. UNIDADE DE ANÁLISE.....	154
IV.7. CARACTERIZAÇÃO E JUSTIFICAÇÃO DAS ÁREAS METODOLÓGICAS ADOPTADAS.....	155
IV.8. A ESTRUTURA METODOLÓGICA DA TESE.....	189
IV.9. OS SISTEMAS INFORMÁTICOS UTILIZADOS NA REALIZAÇÃO DAS ANÁLISES.....	213

CAPÍTULO V. APRESENTAÇÃO E DISCUSSÃO DE RESULTADOS.....	191
V.1. INTRODUÇÃO.....	192
V.2. ANÁLISE DA EVOLUÇÃO DOS INDICADORES AGREGADOS OU “GROSS MEASURES” DE AUDIMETRIA.....	195
V.2.1. Caracterização da evolução das taxas médias de audiências para o total da TV.....	208
V.2.2. Caracterização da evolução das taxas médias de audiências ou <i>Ratings</i>	210
V.2.3. A influência da programação na evolução dos principais indicadores de audimetria..	212
V.2.4. A evolução dos <i>shares</i> de audiências.....	214
V.2.5. A segmentação da emissão diária.....	215
V.2.6. A influência das variáveis sociodemográficas na evolução dos principais indicadores de audimetria.....	216
V.2.7. A análise aos <i>Ratings</i> programáticos: Os modelos explicativos da evolução dos <i>Ratings</i> : A Extração de regularidades	217
V.2.8. A análise dos indicadores comerciais: Os modelos explicativos da evolução dos <i>ratings</i> comerciais a nível absoluto e relativo: A extração de regularidades.....	223
V.3. A ANÁLISE DOS INDICADORES ACUMULADOS DE AUDIMETRIA.....	230
V.3.1. A análise dos “comportamentos estruturais” de exposição televisiva.....	233
V.3.1.1. A análise de fluxos de audiências e de efeitos de herança.....	235
V.3.1.2. A análise da lealdade programática.....	248
V.3.1.3. A análise da lealdade tipológica e ao canal.....	255
V.3.1.4. A interação entre os comportamentos estruturais.....	266
V.3.2. A apresentação de um modelo estrutural da exposição.....	270
V.4. A VIABILIDADE DE CONSTRUÇÃO DOS MODELOS DE EXPOSIÇÃO RECORRENDO AOS MODELOS DE “RATINGS”	280

CAPÍTULO VI. APRESENTAÇÃO DE UMA PROPOSTA PARA O PLANEAMENTO DE <i>MEDIA</i>	317
VI.1. INTRODUÇÃO.....	317
VI.2. ENQUADRAMENTO METODOLÓGICO DA PROPOSTA PARA PROMOVER A CONSTRUÇÃO DOS MODELOS DE EXPOSIÇÃO.....	324
VI.3. OS MODELOS DE EXPOSIÇÃO EM CONTEXTOS DE MAIOR ESTABILIDADE DE VALORES DE <i>RATINGS</i>	332
VI.3.1. As análises univariadas de valores de <i>Ratings</i>	332
VI.3.2. As análises multivariadas de valores de <i>Ratings</i>	337
VI.3.2.1. A análise de séries temporais multivariadas.....	337
VI.3.2.2. A associação de <i>Ratings</i> e de taxas médias de audiências para o total da TV	340
VI.3.2.3. As análises multivariadas de Regressão linear e não linear.....	352
VI.3.3. A comparação de algumas abordagens analíticas de <i>Data Mining</i> e de Estatística Multivariada.....	355
VI.4. A CONSTRUÇÃO DE PLANOS DE <i>MEDIA</i> EM CONTEXTOS DE MAIOR VARIABILIDADE DE VALORES DE <i>RATINGS</i>	355
VI.4.1. Considerações a respeito dos comportamentos de consumo televisivo.....	355
VI.4.2. O conceito de hábito e a segmentação das audiências.....	361
VI.4.3. A selecção de suportes.....	368
VI.4.4. Os modelos de exposição em contextos de existência de hábitos de consumo televisivo.....	371
VI.4.4.1. Os modelos de exposição para uma sequência de suportes quando existe o “Hábito” de ver TV.....	371
VI.4.4.2. Outros contextos de construção de modelos de exposição quando Existe regularidade na exposição.....	387

VI.4.5. Os modelos de exposição em contextos de inexistência de hábitos de consumo televisivo e numa óptica de construção de planos de <i>Media</i>	392
VI.4.5.1. A análise da exposição global.....	399
VI.4.5.2. A análise da análise da exposição aos suportes seleccionados e as teorias explicativas da exposição televisiva.....	411
VI.4.6. A construção do plano final de <i>Media</i>	423
VI.5. OS MODELOS DE EXPOSIÇÃO PARA PROGRAMAS QUE GERAM OBSERVAÇÕES “ATÍPICAS”	428
VI.6. OS MODELOS DE EXPOSIÇÃO QUANDO NOVOS CONTEÚDOS SÃO INTRODUZIDOS NUMA GRELHA DE PROGRAMAS.....	431
 CAPÍTULO VII. CONCLUSÕES E CONSIDERAÇÕES FINAIS.....	437
VII.1. CONCLUSÕES FINAIS.....	437
VII.2. SUMÁRIO DE CONTRIBUIÇÕES.....	443
VII.3. LIMITAÇÕES DAS ANÁLISES REALIZADAS E NOVOS PROBLEMAS DE INVESTIGAÇÃO NA ÁREA DO PLANEAMENTO DE <i>MEDIA</i> E DOS MODELOS DE EXPOSIÇÃO.....	453
 CAPÍTULO VIII. REFERÊNCIAS BIBLIOGRÁFICAS.....	459

APÊNDICES

APÊNDICE A. DESCRIÇÃO E CRÍTICA DOS MÉTODOS DE MEDIÇÃO DAS AUDIÊNCIAS

APÊNDICE B. A INDÚSTRIA TELEVISIVA E A ACTIVIDADE DE PLANEAMENTO DE *MEDIA*

APÊNDICE C. A INVESTIGAÇÃO REALIZADA NO SECTOR TELEVISIVO E ENQUADRAMENTO DA TEMÁTICA PROPOSTA

APÊNDICE D. AS PERSPECTIVAS EXPLICATIVAS DO COMPORTAMENTO DO TELESPECTADOR

APÊNDICE E. PERSPECTIVA DAS ANÁLISES DE PREVISÃO NAS CIÊNCIAS SOCIAIS

APÊNDICE F. APRESENTAÇÃO DE RESULTADOS

APÊNDICE G. COMPARAÇÃO DE METODOLOGIAS

ÍNDICE DE FIGURAS

Figura 1.3.1 – Sequência de análises que viabilizam a apresentação de uma proposta para a previsão das audiências televisivas. Fonte: Elaborado pela autora.....	21
Figura I.5.1 – Estrutura da tese. Fonte: Elaborado pela autora	44
Figura II.3.1 – Factores que condicionam os índices de audiências e a actividade de planeamento de <i>media</i> . Fonte: Elaborado pela autora.....	57
Figura III.1.1.1 – Etapas que integram a construção de uma campanha publicitária. Fonte: Elaborado pela autora.....	70
Figura III.1.3.1 – Etapas que integram a construção de um plano de investimento em publicidade televisiva. Fonte: Elaborado pela autora.....	73
Figura III.1.4.1 – Caracterização de alguns indicadores de audimetria. Fonte: Elaborado pela autora.....	79
Figura III.2.2.1 – Os indicadores que condicionam o investimento em publicidade televisiva. Fonte: Elaborado pela autora.....	85
Figura III.2.3.1 – As metodologias que permitem a construção dos modelos de exposição. Fonte: Elaborado pela autora.....	87
Figura III.5.1.1 – Âmbitos de investigação na área televisiva e disciplinas implicadas. Fonte: Casseti & Chio (1997).....	114
Figura III.5.4.1 – Âmbitos de análise segundo a perspectiva estrutural de análise da exposição televisiva. Fonte: Elaborado pela autora.....	121
Figura IV.1 – O método indutivo e o método dedutivo. Fonte: Elaborado pela autora.....	147
Figura IV.3.1 – Representação do processo de audimetria de medição das audiências televisivas. Fonte: Elaborado pela autora.....	151
Figura IV.4.1 – A construção das bases de dados. Fonte: Elaborado pela autora.....	154

Figura IV.7.1 – Quadro de selecção de metodologias de previsão. Fonte: www.forecastingprinciples.com. Reprinted from J, Scott Armstrong, Principles of Forecasting. Kluwer Publishing, 2001, Revised September 2004R.....	159
Figura IV.7.2 – Árvore metodológica para a selecção de metodologias de previsão. Fonte: www.forecastingprinciples.com.....	160
Figura IV.7.3 – Algumas metodologias de segmentação das bases de dados. Fonte: Elaborado pela autora.....	163
Figura IV.8.1 – Conjunto de metodologias utilizadas na tese. Fonte: Elaborado pela autora.....	167
Figura IV.8.2 – Metodologia conduzida para a pesquisa. Fonte: Elaborado pela autora.....	168
Figura V.3.1 – Análises realizadas aos comportamentos estruturais de exposição televisiva. Fonte: Elaborado pela autora.....	235
Figura V.3.1.4.1 – Fluxograma associado à duplicação das audiências. Fonte: Elaborado pela autora.....	270
Figura V.3.1.4.2 – Contextos de duplicação das audiências. Fonte: Elaborado pela autora.....	271
Figura V.3.2.1 – Proposta de modelo de caracterização do processo decisório de exposição televisiva. Fonte: Elaborado pela autora.....	277
Figura VI.2.1 – Síntese de abordagens que viabilizam o investimento publicitário. Fonte: Elaborado pela autora.....	331
Figura VI.3.2.1 – A previsão das taxas médias de audiências para o total da TV. Fonte: Elaborado pela autora.....	346
Figura VI.3.2.2 – Suportes temporais associados a um programa. Fonte: Elaborado pela autora.....	351
Figura VI.4.2.1 – Segmentos de exposição televisiva do painel de audimetria. Fonte: Elaborado pela autora.....	363
Figura VI.4.3.1 – Identificação das sequências de exposição mais características por telespectador. Fonte: Elaborado pela autora.....	369
Figura VI.4.3.2 – Identificação de segmentos de telespectadores por clustering de sequências de exposição. Fonte: Elaborado pela autora.....	369

Figura VI.4.4.1 – Sequência de análises proposta para a identificação de suportes publicitários. Fonte: Elaborado pela autora.....	372
Figura VI.4.4.1.1 – Sequência de análises para a identificação de suportes. Fonte: Elaborado pela autora.....	373
Figura VI.4.4.1.2 – Duplicação de audiências entre períodos horários. Fonte: Elaborado pela autora.....	385
Figura VI.4.5.1 – Etapas que viabilizam a identificação dos suportes publicitários nos quais investir. Fonte: Elaborado pela autora.....	399
Figura VI.4.5.2.1 – Rede de Bayes para a identificação de associações entre as variáveis que determinam a exposição televisiva. Fonte: Elaborado pela autora.....	419
Figura VI.4.5.2.2 – Base de dados relacional hipotética considerando a informação que consta das bases de audimetria. Fonte: Elaborado pela autora.....	420
Figura VI.4.5.2.3 – Representação de base de dados relacional para a identificação de Regras multirrelacionais. Fonte: Elaborado pela autora.....	421
Figura VI.5.1 – Interação entre o desempenho de diversos suportes horários. Fonte: Elaborado pela autora.....	429

ÍNDICE DE QUADROS

Quadro III.5.3.1 – Diferenciação entre as abordagens qualitativas e quantitativas. Fonte: Espinosa (2006).....	117
Quadro III.6.1 – Enquadramento do estado da arte na área dos modelos de exposição. Fonte: Elaborado pela autora.....	122
Quadro III.6.1.1 – Trabalhos que incidem na previsão dos valores de ratings. Fonte: Exerto do estado da arte apresentado em Danaher et al. (2011) a respeito dos diversos indicadores de audimetria.....	125
Quadro IV.7.1 – Algumas metodologias de análises de Regressão. Fonte: Elaborado pela autora.....	162
Quadro VI.2.1 – Factores que condicionam as análises de previsão e as correspondentes necessidades de análise. Fonte: Elaborado pela autora.....	325
Quadro VI.2.2 – Factores que condicionam as análises de previsão e necessidades de análise. Fonte: Elaborado pela autora.....	326
Quadro VI.3.2.1 – Previsão das taxas médias de audiências para o total da TV. Fonte: Elaborado pela autora.....	350
Quadro VI.3.2.2 – Cálculo de valores de Tvr_Prog% a partir dos dados para períodos horários. Fonte: Elaborado pela autora.....	349
Quadro VI.4.4.1.1 – Selecções programáticas e a avaliação das influência das variáveis estruturais. Fonte: Elaborado pela autora.....	370
Quadro VI.4.5.1.1 – Identificação dos suportes por telespectador e dia e construção das sequências mais representativas. Fonte: Elaborado pela autora.....	401
Quadro VI.4.5.1.2 – Identificação de sequências de exposição. Fonte: Elaborado pela autora.....	406
Quadro VI.4.5.1.3 – Identificação de suportes por telespectador. Fonte: Elaborado pela autora.....	409
Quadro VI.4.5.1.4 – A construção de sequências transformadas. Fonte: Elaborado pela autora.....	410

Quadro VI.4.5.1.5 – Identificação de sequências com cobertura mínima. Fonte: Elaborado pela autora.....	410
Quadro VI.4.5.2.1 – Algumas variáveis que determinam a exposição televisiva e a selecção dos conteúdos. Fonte: Elaborado pela autora.....	412
Quadro VI.4.6.1 – Identificação dos suportes a incluir nas grelhas de exposição onde incidem os algoritmos de optimização. Fonte elaborado pela autora.....	421

ÍNDICE DE TABELAS

Tabela V.2.9.1 – Coeficientes de determinação dos modelos de Regressão linear. Fonte: Elaborado pela autora.....	228

LISTA DE ABREVIATURAS, SIGLAS, SÍMBOLOS OU ACRÓNIMOS

ABREVIATURAS

AT – Antes espaço informativo da noite

ATV – *Average Time Viewing* (Indicador de audimetria que representa o tempo médio de exposição a um suporte, independentemente de ser considerado um telespectador em função do tempo de exposição mínimo definido para ser considerado como tal)

ATS – *Average Time Spent* (Indicador de audimetria que representa o tempo médio que um telespectador despendeu a ver um evento ou canal)

BP – *Back Propagation* (Algoritmo para treino de Redes Neurais MLP)

C.C. – Caso Contrário

COB. – Cobertura (Percentagem do grupo-alvo que contacta, pelo menos uma vez, com um suporte ou combinação de suportes)

CP – Componentes Principais na análise Factorial

D. Casa – Dona de Casa (Variável de segmentação das bases de dados de audimetria)

DIF – DIFerencial

DIM – DIMensão associada a um plano factorial na análise MDS

DM – *Data Mining*

DMT – *Data Mining* Temporal

DT – Depois espaço informativo da noite

FC – Futebol Clube

FID – FIDelidade programática (Indicador de audimetria que quantifica a fidelidade programática a um suporte)

FS – Fim-de-Semana

GO – Grupos Ocupacionais (Variável de segmentação que permite a identificação das classes sociais representadas nas bases de dados de audimetria)

G&U – Gratificações e Usos do *Media* (Uma perspectiva explicativa da exposição televisiva)

GRP – *Gross Rating Points* (Indicador de audimetria que quantifica as audiências acumuladas de um plano de *media*. Medida comparativa do impacto quantitativo de uma campanha publicitária. Corresponde à soma das audiências (*ratings*) de todos os suportes de um plano de meios (um GRP é a probabilidade de contactar com 1% do grupo alvo))

H – Hora

IA – Índice de Apreciação

IC – Intervalo de Confiança

LC – Litoral Centro (Uma região associada à variável de segmentação das bases de dados de audimetria referente às regiões geográficas)

LDIN e LDOUT – Programa de LeaD_IN (antecessor) e programa de LeaD_OUT (sucessor) ao programa em análise

LN – Litoral Norte que representa uma região nas bases de dados de audimetria

LP – Linear Programming

ND Casa – Não Dona de Casa

PT – Prime Time (Designa o horário nobre que corresponde a uma das faixas horárias em que se subdivide a emissão diária atendendo à classificação adoptada pela indústria televisiva)

RAT% – *Rating* em percentagem. Indicador de audimetria que representa a audiência média de um suporte. Quando surge com a designação de *RAT_Prog%* diz respeito a um suporte horário, *RAT_Periodo%* diz respeito a um suporte horário e *RAT_Spot%* diz respeito a uma inserção comercial

RLM – Regressão Linear Múltipla

RL – Regressão Linear

SHR% ou SHARES – Indicador de audimetria que representa a quota de audiência de um evento, isto é, de um canal (*Shr_Período%*) ou de um programa (*Shr_Prog%*). Relação entre o tempo total despendido a ver um determinado canal/suporte/programa, relativamente ao tempo total despendido a ver televisão, num determinado período de tempo, por exemplo, hora, dia, semana, mês.

S – Semana

SEG. – Segundos

Spot_Ratio% – Indicador calculado e que resulta do quociente entre os valores dos *ratings* programáticos e comerciais

ST – Séries Temporais

SYN – Sinergias entre tipologias programáticas dos programas, isto é, a coincidência das tipologias

TS – *Time Series*

TV – TeleVisão

TVR% – Indicador de audimetria designado por taxas média de audiências para o total da TV. Quando surge com a designação de *TVR_Prog%* diz respeito a um suporte horário e a designação *TVR_Periodo%* diz respeito a um suporte horário

Vs. – VerSos

ACRÓNIMOS

AID – Automatic Interaction Detector (Metodologia de *Data Mining* de Árvores de Regressão)

AFC – Análise Factorial de Correspondências

AFCP ou ACP – Análise Factorial por Componentes Principais ou Análise em Componentes Principais

ALEA – Acção Local Estatística Aplicada

ANACOM – Autoridade Nacional de COMUNICAções

ANOVA – ANalysis Of VAriance (Análise de Variância)

APRIORI – Metodologia de *Data Mining* destinada à análise de bases de dados transacionais e que viabiliza a construção de Regras de Associação

AR – AutoRegressive

ARMA – AutoRegressive Moving Average (Metodologias AutoRegressivas e Médias Móveis de previsão de séries temporais)

ARIMA – AutoRegressive Integrated Moving Average (Metodologias Autorregressivas Integradas de Médias Móveis de previsão de séries temporais)

BBC – *British Broadcasting Channel*

BSO – Boolean Symbolic Object (Objecto Simbólico Booleano)

C5 – Algoritmo de *Data Mining* que corresponde a uma Árvore de Classificação

CART – Metodologia de *Data Mining* Proposicional que corresponde a uma tipologia de Árvore de Classificação e Regressão

CHAID – CHi-square Automatic Interaction Detector (Metodologia de *Data Mining* Proposicional que corresponde a uma tipologia de Árvore de Classificação e Regressão)

CLARA – CLustering LARge Applications (Algoritmo de *clustering* para grandes bases de dados)

CPCA – Center Principal Component Analysis (Algoritmo de análise ACP simbólica que corresponde ao método dos centros)

CPM – DeMographic Cost per thousand

CPD – Conditional Probability Distribution (Em Redes Bayesianas representa a probabilidade de cada estado da variável atendendo às combinações totais dos estados dos nós dos quais depende)

CPU – Processador de um Computador

COWEB – Algoritmo de *Data Mining* de *Clustering*

CRM – Customer Relationship Management

CRECS – Conselho Regulador da Entidade Reguladora para a Comunicação Social

CTA – Algoritmo de análise ACP simbólica

DBCSAN – Algoritmo de *Data Mining* de *Clustering*

DENCLUE – Algoritmo de *Clustering* de *Data Mining*

DISS – DISSimilarity measures for symbolic data (Medidas de dissimilaridade para bases de dados simbólicas)

DIV – DIVisive *clustering* (Metodologia de *Clustering* Divisivo de bases de dados simbólicas)

DFT – Discrete Fourier Transform

DNA – Deoxyribonucleic Acid

DTW – Dynamic Time Wrapping (Metodologia para alinhamento de séries temporais)

DWT – Discrete Wavelet Transform

EAM – Erro Absoluto Médio

ERC – Entidade Reguladora para a Comunicação Social

GFK – Empresa de medição das audiências televisivas

GLM – General linear Models

GM – Graph Mining

GPS – Global Positioning System

HMM – Hidden Markov Models

HYPR – Hierarchical and Pyramidal *Clustering* of Symbolic data

IBL – Instance Based Learners

IBM – Instance Based Methods

ICA – Independent Component Analysis

ILP – Inductive Logic Programming

INE – Instituto Nacional de Estatística

ISCTE - Instituto Superior de Ciências do Trabalho e da Empresa

IST – Instituto Superior Técnico

JAR – Journal of Advertising Research

KDD – Knowledge Discovery and *Data Mining*

KMEDOIDS – Partitioning technique of *clustering*

KNN – K Nearest Neighbours (Classificador de dados baseado no vizinho mais próximo)

LSD – Teste utilizado na Análise de Variância

MARKTEST – Empresa que realiza a medição das audiências televisivas segundo o processo de audimetria

MBR – Memory Based Reasoning

MDS – MultiDimensional Scaling (Metodologia de análise Estatística Multivariada de Escalonamento Multidimensional)

MIX-PCA – Algoritmo de análise factorial em Componentes Principais simbólico

MLP – MultiLayer Perceptron Neural Networks (Redes Neurais Perceptrão Multicapa)

MMDR – Machine Methods for Discovering Regularities

MRDM – Multirelational *Data Mining* (*Data Mining* Multirrelacional)

MRDTL – Multi-Relational Decision Tree Learning

MR SMOTI – Algoritmo de Árvores de Regressão Relacionais

MSE – Mean Square Error (Erro quadrático médio)

MS – Mean Square

NN – Nearest Neighbours (Observação mais próxima da observação a prever ou a classificar)

OPTICS – Algoritmo de *Clustering*

PAM – Partitioning Around Medoids (Algoritmo de *Clustering*)

PCA – Principal Component Analysis (Metodologia de Estatística Multivariada de Análise Factorial em Componentes Principais)

PES – Probabilistic Expert Systems (Os sistemas expertos probabilísticos)

PLS – Partial Least Squares Regression (Regressão de Mínimos Quadrados Parciais)

PUBLITEXT – Base de dados de audimetria de audiências comerciais

PVR – Personal Video Recorder

RBF – Radial Basis Functions (Redes Neurais de Funções de Base Radial)

RMSE – Relative Mean Square Error (Erro Quadrático Médio Relativo)

RMS – Root Mean Square (Erro Quadrático Médio)

RMLF – Rendimento Mensal Liquido do agregado familiar

RT_SPCA – Algoritmo de análise factorial simbólico

RT_PCA – Algoritmo de análise factorial simbólico

SARIMA – Metodologia de Previsão de Séries Temporais

SAS – *Software* de Análise Estatística de Dados

SCLUST – Algoritmo de *Clustering* de bases de dados simbólicas

SDT – Symbolic Decision Trees

SMO – Sequential Minimal Optimization

SMOTI – Algoritmo de Árvore de Regressão que analisa dados proposicionais

SO – Symbolic Objects (Objectos Simbólicos)

SODAS – *Software* para a análise de bases de dados simbólicas

SOM – Redes Neurais de *Clustering* SOM (Kohonen Neural Networks)

SGCA – Symbolic Generalized Canonical Analysis

SPCA – Symbolic Principal Component Analysis (Análise PCA simbólica)

SQL – Structured Query Language (Linguagem para o manuseamento de bases de dados)

SSDM – Semi Structured *Data Mining* (Data Mining Semi Estructurado)

STATISTICA – *Software* de Análise Estatística de Dados

STEPWISE – Regressão Linear Passo a Passo

STOCHAPPROX – Algoritmo de análise factorial simbólico

SVD – Singular Value Decomposition (Método de factorização)

SVM – Support Vectors Machines (Máquinas de Suporte Vectorial)

SWOT – Metodologia para a análise interna e externa da posição competitiva de uma empresa

SYKSOM – Symbolic SOM Neural Networks (designa um algoritmo de Redes SOM simbólicas)

TDM – Temporal *Data Mining* (*Data Mining* Temporal)

TDT – Televisão Digital Terrestre

TELEREPORT – Base de dados de audimetria de audiências programáticas

TILDE – Designa um algoritmo que permite o desenvolvimento de uma Árvore em contexto de ILP (“Top-down Induction of First-order Logical Decision Trees”)

TLC – Teorema do Limite Central

TREE – Metodologia de *Clustering* simbólica

TSDB – Time Series Database (Base de Dados Temporal)

TSDM – Time Series *Data Mining* (*Data Mining* de Séries Temporais)

UER – La Unión Europea de Radiodifusión

USA – United States of America (Estados Unidos)

VC – Vapnik Chervonenkis

VIF – Variance Inflator Factor

VOD – Video On Demand

VPCA – Vertices Principal Component Analysis (Análise Factorial Simbólica de vértices PCA)

VQ – Vector Quantization (Algoritmo de *Clustering*)

WARC – World Advertising Research Center

XML – Extensible Markup Language

GLOSSÁRIO

AUDIMETRIA – Tecnologia de medição de audiências que permite a recolha de dados que caracterizam a exposição televisiva e o cálculo dos índices de audiências programáticos e comerciais.

BREAK – Designa um espaço comercial no *software* Telereport e Publitex. Representa um conjunto de anúncios que se emitem de forma agrupada, com regras definidas em termos legais ou com normas definidas pelas estações emissoras, identificado por “separadores” que delimitam o seu início e fim dentro de um programa ou entre programas.

CLASSE – Na Análise Simbólica representa um conjunto de objectos que partilham os mesmos atributos, operações, métodos, relacionamentos e semântica.

CLUSTER – Conjunto de observações que deriva de algoritmos de análise de agrupamento de dados.

CLUTTER – Representa um fenómeno complexo ao nível da percepção e eficácia da mensagem publicitária, que se encontra associado com a duração das inserções comerciais e dos espaços comerciais.

CLASSE DE OBJECTOS – Na Análise Simbólica representa um conjunto de ideias, abstrações ou coisas do mundo real, bem delimitados e definidos, cujas propriedades e comportamento seguem as mesmas Regras.

CONCEITO – Unidade de pensamento formada através de abstrações baseadas em características comuns a um conjunto de objectos.

CO -VIEWING – Exposição programática influenciada pelo grupo em que o telespectador se encontra inserido.

HAMMOCKING – Estratégia de programação que consiste em transmitir um programa de fraca audiência entre dois programas mais populares, por forma a maximizar a sua exposição televisiva.

LEAD IN (ou Ldin) e LEAD OUT (Ldout) – Programa que antecede (“lead in”) ou sucede (“lead out”) o programa de referência.

PAINEL – Caso particular da amostra longitudinal em que a totalidade das unidades amostrais se mantêm ao longo do tempo de vida da amostra.

RATING – Representa um indicador de audimetria que condiciona maioritariamente a venda do espaço comercial e os conteúdos a transmitir e que também pode ser designado na bibliografia da especialidade por taxas médias de audiências.

“RATINGS” – Designa em sentido lato as bases de dados que são obtidas recorrendo ao processo de audimetria de medição das audiências televisivas.

RAW DATA – Dados de audimetria que quantificam a exposição televisiva a nível individual, podendo ser, também, designados de dados dos ficheiros base de audimetria.

REPEAT VIEWING – Designação associada à duplicação de audiências entre episódios ou transmissões de um mesmo programa ou faixa horária no decorrer do tempo e que quantifica o grau de lealdade programática.

ROADBLOCKING – Transmissão simultânea de uma inserção comercial em várias estações televisivas.

SPOT – Designação de inserção publicitária no *software* Publitex.

SPLITS – Em Árvores de Decisão designa o processo de fraccionar a amostra de dados em dois ou mais subconjuntos.

SURROGATE – Em Árvores de Decisão designa uma variável substituta quando se verifica o *split*, isto é, a partição de um nó.

TARGET – Segmento de audiência-alvo.

ZAPPING – Comportamento de mudança de canal realizado no decorrer dos espaços comerciais, afim de evitar a publicidade. “Zappers” designa os telespectadores que realizam o *zapping*.

ZIPING – Comportamento dos telespectadores em que se pretende evitar os espaços comerciais quando assistem a um vídeo, realizando o seu “fast-forwarding”.

CAPÍTULO I

INTRODUÇÃO

Neste capítulo, descreve-se o contexto em que surge a temática a desenvolver – a viabilidade dos modelos de exposição¹, que servem de suporte à actividade de planeamento de *media*² e à selecção e estruturação dos conteúdos a incluir nas grelhas de programação. A audimetria corresponde a um instrumento de medição das audiências televisivas, que permite uma análise diacrónica, rápida e mais detalhada do consumo televisivo. Num domínio classificado com imprevisível e indeterminado, constitui um instrumento de previsão e planeamento. A inferência indutiva depende da regularidade dos comportamentos dos telespectadores e dos indicadores que deles resultam. Nesse sentido, as análises exploratórias, descritivas e explicativas são necessárias para uma definição dos modelos quantitativos de *media* mais adequados á evolução dos dados. Pretende-se analisar a “exposição televisiva”, recorrendo a análises estruturais³ explicativas da evolução dos “ratings”⁴, entre as quais a da duplicação das audiências⁵ entre os suportes publicitários.

¹ Os modelos de exposição incidem na previsão da cobertura e frequência que um conjunto de suportes poderá alcançar (*Craig & Ghosh, 1993*), por análise da duplicação das audiências intra e inter-suportes ou, em perspectivas alternativas, de análise que permitam avaliar as audiências acumuladas de um plano ou contribuir para a selecção dos suportes a considerar nas etapas posteriores das análises que visam a construção de uma campanha publicitária. Em suma, caracterizam e quantificam a exposição televisiva.

² A atividade de *planeamento de media* consiste na construção de planos de investimento nos *media*. Os planos devem maximizar a exposição ao público-alvo mas com frequências consideradas “óptimas”, e tendo em atenção determinadas restrições orçamentais (*McMellon & Eftimov, 2004*.) ou relacionadas com as características dos suportes publicitários, ou seja, dos programas ou tempos aos quais associar as inserções comerciais.

³ As “análises estruturais” correspondem a uma perspectiva de caracterização e avaliação dos factores que condicionam a exposição televisiva, como, por exemplo, a influência da estruturação da grelha e a análise dos “comportamentos estruturais” que decorrem dos factores de planeamento, entre os quais a “lealdade”.

⁴ Os “ratings” designam os dados obtidos pelo processo de audimetria de medição das audiências televisivas.

⁵ A *duplicação de audiências* engloba diversas perspectivas de análise que avaliam a tendência que se verifica nos telespectadores para estarem presente nas audiências partilhadas por diversos suportes publicitários.

I.1. A PROBLEMÁTICA DA INVESTIGAÇÃO: OS MODELOS DE EXPOSIÇÃO NA ACTIVIDADE DE PLANEAMENTO DE MEDIA

A televisão corresponde ao meio audiovisual que assume uma maior preponderância no preenchimento do tempo livre dos cidadãos, mantendo, na actualidade, a sua liderança ao nível do investimento publicitário. Por conseguinte, este meio de comunicação tem sido objecto de análise em inúmeras áreas e segundo diversas perspectivas, despertando um crescente interesse em múltiplas organizações académicas e empresariais.

No sector televisivo, facilmente identificamos uma grande diversidade de temáticas, disciplinas, instrumentos e áreas de investigação (*Cassetti & Chio, 1997*). Podemos diferenciá-los em trabalhos de investigação qualitativa ou quantitativa (*Stavitsky, 2000*), empresariais ou académicos, e associados a diversas áreas científicas, como, por exemplo, à Economia, Gestão, Psicologia, Marketing e Sociologia. Cada área aborda questões diferenciadas, recorrendo a distintas fontes de dados e métodos de investigação. Efectivamente, de acordo com *Hartley (2000, p. 34, cit. in Huertas, 2002, p. 54)*, “La televisión es demasiado grande como sistema textual, demasiado complicada en todas sus facetas de producción, programación y recepción, demasiado variada en tiempos y lugares como para ser estudiada como una entidad única.”

Na opinião de *Huertas (2002)*, a justaposição das diferentes perspectivas de análise conduz a um problema de integração, não existindo uma recopilação sistemática que solucione as inúmeras dificuldades que decorrem da fragmentação do conhecimento das audiências televisivas. A autora salienta a dificuldade de desenvolvimento de um referencial teórico que articule o conjunto de estudos empíricos realizados no decorrer do último século, como consequência da complexa natureza deste meio audiovisual.

A presente tese procura analisar os modelos matemáticos de previsão da exposição televisiva, os quais permitem ao sector tentar conhecer antecipadamente os níveis de audiência dos suportes publicitários nos quais pretende investir.

A relevância desta temática advém das inúmeras limitações que apresentam os diversos modelos de *media* actualmente existentes e que são responsáveis pelo elevado desperdício de recursos financeiros que afecta múltiplas organizações que, de alguma forma, dependem da publicidade televisiva (Makienko, 2012). Os modelos de *media* podem ser classificados em modelos de exposição, modelos de optimização e modelos de avaliação dos efeitos da exposição à publicidade nas vendas e nos comportamentos dos consumidores. Os modelos de exposição incidem na compreensão da evolução das audiências televisivas e na sua posterior previsão. Apesar da sua relevância para a gestão das organizações, estes modelos não têm sido privilegiados na investigação académica.

Várias disciplinas podem e devem contribuir para a construção dos modelos de *media*, destacando-se as Ciências Matemáticas e o Marketing, assim como todas as restantes que viabilizem um maior conhecimento sobre os factores que determinam o consumo televisivo e a respeito das perspectivas de análise que podem conduzir aos melhores resultados quando é necessário realizar análises de previsão nas Ciências Sociais. Esta temática apresenta, inevitavelmente, uma grande abrangência ao nível dos conceitos teóricos e metodológicos que lhe são subjacentes, e sem os quais não será viável alcançar uma solução solidamente fundamentada e exequível, que permita contribuir para que sejam ultrapassadas as inúmeras dificuldades que estes modelos colocam ao profissionais do sector.

Numa tentativa de resumir a vasta diversidade de âmbitos de análise existentes na área televisiva e, assim, poder enquadrar o objectivo proposto neste trabalho, verificamos que, segundo Casseti e Chio (1997), existem cinco orientações teóricas, que se diferenciam em função do objecto central do estudo a desenvolver: o *media*, o texto, o telespectador, os contextos e a recepção. A presente tese incide na análise do “telespectador” e neste contexto insere-se, segundo os autores, a análise dos índices de audiências, do número e da identidade dos telespectadores. Atribuir uma maior ênfase a determinado elemento implica que sejam realizadas certas escolhas práticas quanto às disciplinas, instrumentos e perspectivas de análise que devem ser adoptadas.

Os modelos de exposição visam contribuir para uma maior rentabilização do investimento publicitário e, logo, incidem na análise dos dados que condicionam a venda do espaço e tempo comercial e a definição de uma estratégia de comunicação mais adequada ao segmento-alvo. Por conseguinte, torna-se necessário o desenvolvimento de modelos estruturais explicativos e/ou modelos de previsão das audiências televisivas, recorrendo à investigação quantitativa e a fontes de dados designadas na bibliografia da especialidade por “*ratings*” (Stavistky, 2000). A inevitabilidade do recurso aos dados de audimetria quando se pretende a construção dos modelos de exposição está demonstrada na bibliografia da especialidade (Zarate, 1997).

Em Casseti e Chio (1997), são, também, identificadas onze áreas de análise onde convergem determinadas disciplinas, instrumentos e metodologias de investigação. Uma das áreas apresenta, como objecto de estudo, a análise das entidades e da estrutura da audiência, ou seja, incide na análise dos comportamentos de consumo televisivo, com o intuito de captar a sua dimensão quantitativa, numa óptica explicativa ou de previsão. Neste contexto, é avaliada a dimensão da recepção televisiva, descrevendo, também, alguns dos seus aspectos estruturais, como, por exemplo, a distribuição do consumo no decorrer do tempo. A evolução dos dados que quantificam a exposição televisiva é avaliada ao nível da influência das variáveis de planeamento, do acesso ao *media* ou da disponibilidade dos telespectadores, entre outras variáveis essencialmente estruturais.

Esta vertente de análise introduziu na metodologia de investigação os instrumentos de carácter quantitativo, numérico e estatístico que têm como função registar os comportamentos dos telespectadores com o maior rigor e de forma sistemática, ou seja, a audimetria (Jauset, 2008) e os diários de audiências, apresentando estes últimos maiores limitações (Mariet, 1990). A investigação quantitativa das audiências e a quantificação da exposição televisiva correspondem a áreas que não têm sido privilegiadas no passado. De acordo com Huertas (2002, p.151), “En la actualidad, aunque cada vez son más los investigadores interesados en la cuantificación de la audiencia, todavía es muy reducida la bibliografía especializada e, incluso, la que dedica aunque sólo sea un apartado a este asunto.”

A construção de um sistema que fosse credível ao nível da investigação quantitativa das audiências representou uma das metas mais relevantes no âmbito da investigação referente a este sector de actividade (*Casetti & Chio, 1997*). Porém, a vertente quantitativa de análise das audiências encontra-se muito pouco desenvolvida, nomeadamente ao nível dos métodos analíticos adoptados (*Weber, 2002*), assim como relativamente ao desenho das amostras que são utilizadas nos processos de sondagens das audiências televisivas (*Huertas, 2002*).

Actualmente, a televisão (TV) permanece, a nível mundial, como o *media* que maximiza o investimento realizado em publicidade (*Callan, 2009; Yao, 2009; ZenithOptimedia, 2009*). Um estudo de audiência, na perspectiva do planeamento de *media*, incide na análise da exposição aos *media* e na sua quantificação e previsão, constituindo toda a informação de base necessária ao trabalho que visa a definição de uma estratégia de comunicação. Estas análises condicionam a selecção da melhor solução, do ponto de vista estratégico e tático, de aquisição de espaço e tempo comercial, e sempre de acordo com os objectivos de Marketing e o orçamento publicitário de um anunciante.

Actualmente, é necessário que as campanhas publicitárias sejam construídas da forma mais eficiente possível, devido às estratégias cada vez mais agressivas que podem ser observadas entre as estações televisivas, aos reduzidos recursos de que dispõem os anunciantes e a uma crescente oferta ao nível dos *media* (*Makienko, 2012*). A construção de um plano de investimento em publicidade televisiva implica a selecção dos *media*, suportes, formatos e número de inserções mais adequado aos objectivos e limitações que determinada campanha publicitária apresenta. O tempo adquirido em intervalos publicitários é função da previsão de alguns índices de audiências ou da exposição do segmento-alvo e os custos que lhe estão associados apresentam uma tendência crescente (*The Ad Class, 2010*). Inevitavelmente, verificamos que existe uma crescente necessidade ao nível do desenvolvimento de modelos explicativos e de previsão de múltiplos indicadores de audimetria (*Givon & Grosfeld-Nir, 2008; Kelton & Schneider-Stone, 1998; Weber, 2002*).

Os indicadores de audimetria podem ser classificados em indicadores acumulados e agregados, destacando-se os valores de *ratings*, que podem, também, ser designados de taxas médias de audiências. Os *ratings* condicionam a venda do tempo comercial no “spot market”, onde é possível adquirir um menor conjunto de inserções publicitárias, assim como quando o investimento é realizado no “up-front market” e no “scatter market”, recorrendo à construção de planos de *media*, uma vez que, do seu somatório, resulta o valor das audiências acumuladas do plano.

Os índices de audiência que constam das bases de dados de audimetria, por sua vez, resultam do comportamento dos telespectadores que integram um painel de análise da exposição televisiva. As análises de previsão dos comportamentos de consumo televisivo, ou de qualquer outro comportamento, apresentam inúmeras limitações. As Ciências Sociais são classificadas como inexactas e, logo, a sua descrição é, sobretudo, qualificativa. Por conseguinte, nesta área, o conceito de análise de previsão aproxima-se, por vezes, em maior grau do explicativo, uma vez que, frequentemente, apenas determina os factores que influenciam a propensão para se verificarem determinados resultados no decorrer do tempo. Quer dizer, esta análise assume características que a diferenciam das análises que podem ser realizadas nas Ciências Exactas e duras, onde previsões para instantes concretos no espaço e no tempo podem apresentar um considerável grau de precisão.

Hoje em dia, os índices de audiência permitem a avaliação do desempenho das organizações, para além de constituírem um meio de diagnóstico, planificação e apoio à gestão das estações televisivas e de todos os restantes agentes intervenientes neste sector. O planeamento de *media* apenas será viável recorrendo a determinados modelos matemáticos, quer ao nível da optimização, quer considerando as análises quantitativas de previsão que incidem na avaliação da cobertura e frequência de um plano de *media*, nos valores de *ratings* ou na selecção dos suportes mais adequados a determinada campanha publicitária. Numa fase posterior, será necessário o acompanhamento e a avaliação dos seus resultados, os quais podem ser quantificados através dos indicadores de audimetria, sendo o mais relevante os *ratings*.

O custo da publicidade é função dos valores de *ratings*, que, por sua vez, condicionam os valores das audiências acumuladas previamente definidos como metas a alcançar no decorrer do horizonte temporal em que se desenrola uma campanha publicitária. Uma maior percepção sobre os factores que condicionam a evolução dos *ratings* e a viabilidade de realizar previsões com reduzidas margens de erro é essencial, uma vez que pequenos desvios podem originar grandes perdas ou poupanças aos anunciantes e às estações televisivas. A este propósito, *Danaher, Dagger e Smith (2011. p.1)* afirmam: “We demonstrate that making improvements in *ratings* forecasts could save the television industry between \$250 and \$586 million per year”. A necessidade de conhecer a estrutura dos comportamentos de consumo televisivo, a evolução dos fluxos de audiências entre os suportes publicitários e dos *ratings* que deles resultam, e de dispôr de dados objectivos sobre onde, como e quanto investir no *media* e nos suportes publicitários é fundamental para uma maior rentabilização do investimento financeiro realizado pelos anunciantes.

Nesta área, as melhores perspectivas para a construção dos modelos de exposição dependem, também, marcadamente dos contextos e, conseqüentemente, os resultados não são generalizáveis quando estes diferem. A investigação realizada no sector televisivo apresenta, frequentemente, conclusões contraditórias, inclusive quando nos restringimos a um mesmo âmbito de análise. O factor temporal, cultural, socioeconómico e de estruturação do mercado televisivo influenciam fortemente os resultados, assim como os métodos de investigação e as fontes de dados adoptadas.

A construção dos planos de *media*, após o desenvolvimento da audimetria, é, essencialmente, avaliada por análise de alguns indicadores de audimetria, sendo difícil conseguir encontrar abordagens alternativas. A bibliografia nacional identificada na área dos modelos de *ratings* é muito reduzida, incidindo em dados e em perspectivas de análise que não correspondem às mais relevantes para o planeamento de *media*, e, por outro lado, em períodos que não têm em atenção as consideráveis alterações que se têm verificado neste sector no decorrer da última década.

No âmbito da previsão dos índices de audiências, a investigação internacional mais recente apresenta, ainda, consideráveis margens de erro e, de acordo com *Danaher et al. (2011)*, são muito escassas as publicações existentes, sendo inúmeras as suas limitações. A análise dos erros de previsão consta de um reduzido número de publicações (*Napoli, 2001; Yoo & Kim, 2002*) que descrevem resultados com reduzidas percentagens de variância explicada, assim como apenas pontualmente se tem avaliado a repercussão do nível de agregação dos dados na dimensão dos erros de previsão (*Meyer & Hyndman, 2005*).

Os modelos explicativos da evolução dos *ratings* e de outros indicadores agregados e acumulados de audimetria, nomeadamente dos valores de *shares* e da exposição global ao *media* ou considerando os comportamentos designados de “estruturais”, são, também, escassos na bibliografia internacional (*Stavitsky, 2000*), para além de imperar a utilização de metodologias mais tradicionais de análise estatística dos dados. A mesma tendência é observada ao nível do desenvolvimento de modelos integrados explicativos do comportamento dos telespectadores.

Em suma, persistem inúmeras lacunas na bibliografia internacional e nacional relacionada com os modelos de exposição, a diversos níveis. A maturidade de uma área, ou seja, o conhecimento que detemos a respeito do funcionamento de sistemas que são considerados complexos, apresenta os fundamentos da sua previsão, ao nível da sua viabilidade e da identificação das perspectivas que podem conduzir aos melhores resultados.

O principal objectivo deste trabalho reside em contribuir para o desenvolvimento dos modelos de *ratings* e, em especial, para a identificação das melhores perspectivas de construção dos modelos de exposição. O estudo incide, inevitavelmente, na temática da análise estrutural da “exposição televisiva”, através do desenvolvimento de modelos de *ratings*, e na avaliação da viabilidade da sua previsão em determinados contextos de análise que condicionam maioritariamente o investimento publicitário.

I.2. A RELEVÂNCIA DO TEMA DE INVESTIGAÇÃO

O desenvolvimento de modelos de *ratings* determina o desempenho e a sobrevivência das inúmeras organizações ligadas a este sector de actividade. De acordo com *Fernandes (2001)*, o funcionamento estrutural da televisão comercial é função do que assegura a viabilidade económica das empresas de comunicação social, ou seja, da publicidade. A autora refere, também, que o poder económico impôs uma nova lógica à programação televisiva, a qual assenta em grelhas de programação que permitam gerar os maiores níveis de audiência e, logo, de receitas. Por conseguinte, a grelha de programas representa, actualmente, um instrumento de acção estratégica das estações de televisão.

A selecção e estruturação dos conteúdos transmitidos visa a fidelização do público e apoia-se, quer na análise dos índices de audiências, que quantificam a exposição ao *media* sob diversas vertentes, quer em estudos de audiência que caracterizam e explicam a estrutura dos comportamentos de consumo televisivo. Quer dizer, a grelha de programas dos principais canais generalistas obedece ao critério económico e aos estudos de audiência. Segundo *Fernandes (2001)*, planejar e organizar uma grelha de programação corresponde a uma tarefa anual, semanal e diária. Os modelos de *ratings* são essenciais em todo este processo. A autora salienta, por exemplo, que, anualmente, tem que ser decidido qual será o valor da verba a disponibilizar para a aquisição dos diversos conteúdos programáticos. O custo dos conteúdos é função da exposição histórica alcançada e do seu valor futuro, quantificado através dos inúmeros indicadores que constam das bases de audimetria. A mesma autora refere, ainda, que, semanalmente e diariamente, tenta promover-se a fidelização do público, variando a grelha consoante a exigência dos anunciantes e a informação que consta dos estudos de audiência. O grau de lealdade dos telespectadores e a dimensão da duplicação das audiências entre os suportes publicitários são algumas das variáveis que influenciam as alterações a realizar. Apenas com a audimetria será viável a quantificação dos níveis de audiência conjunta dos suportes publicitários, a nível temporal ou programático.

Na perspectiva do anunciante, a actividade de planeamento de *media* incide na identificação da melhor solução estratégica e tática de aquisição do espaço comercial. Neste contexto, o que se denomina de investigação ou estudos de audiência considera, como um dos seus objectos de estudo, a informação necessária para seleccionar de forma óptima os espaços e tempos publicitários, viabilizando, assim, que a mensagem apresente uma considerável cobertura do segmento-alvo a determinados níveis de frequências de exposição e ao menor custo possível (*Espinosa, 2006*).

O planeamento só poderá ser realizado dispondo de informação contínua, rápida, detalhada e fiável sobre os comportamentos de consumo televisivo. A audimetria representa um instrumento de medição das audiências que apresenta as maiores vantagens, assim como qualquer outra tecnologia actual ou futura que lhe seja equivalente ao nível das características dos dados que produza, como, por exemplo, o “*audio-matching*”. Por conseguinte, os modelos explicativos e de previsão de *ratings* são essenciais para o planeamento de *media*.

Uma vez que a actividade de programar em televisão pretende produzir público para os anunciantes, é necessário compreender como os telespectadores interagem com o *media* e, posteriormente, conseguir realizar a previsão da exposição aos suportes publicitários, ou seja, a programas, anúncios ou faixas horárias. Na opinião de *Zarate (1997)*, a melhor perspectiva quando se pretende realizar uma previsão em Ciências Sociais corresponde ao conhecimento da estrutura dos comportamentos, uma vez que deverá apresentar uma menor variabilidade no curto prazo e na quantificação da propensão para se verificarem determinados resultados. Quer dizer, as análises explicativas e de previsão condicionam-se reciprocamente e apenas a audimetria tem, maioritariamente, viabilizado, no passado, a análise de comportamentos que são irregulares e que não apresentam ciclos com uma curta duração. De acordo com *Chakrapani (1993)*, alguns autores defendem que não é necessário explicar um fenómeno, desde que o comportamento possa ser previsto de forma precisa, o que não se verifica nesta área. O autor salienta que, na realidade, seria desejável que os modelos apresentassem capacidade explicativa dos resultados que produzem.

Também para *Patel e Schlijper (2008)*, o grau de lealdade que os consumidores manifestam condiciona os modelos de previsão que devem ser adoptados. Os autores salientam que, em contextos de maior lealdade, um sistema determinista e contínuo corresponderá a um sistema de equações ordinárias diferenciais. Quando não existe lealdade, os mesmos autores constataram que o valor da quota de mercado de um produto não apresenta um comportamento dinâmico e dependerá, unicamente, do que está disponível em determinado momento, o que dificulta as análises de previsão.

Por conseguinte, os comportamentos estruturais de lealdade devem ser avaliados e construído um modelo explicativo da exposição televisiva, que, por sua vez, condicionará as melhores perspectivas de previsão. A estrutura da programação, os canais e o mercado condicionam fortemente o consumo televisivo e são responsáveis pela ausência de padrões relacionados com a manifestação de reais preferências e interesses do público. Quanto maiores forem as influências estruturais, mais abrangente e complexo poderá ser o modelo explicativo dos seus comportamentos, mas, do ponto de vista da previsão, é provável que possam vir a ser alcançados melhores resultados, desde que exista uma exposição mais regular ao *media*, a qual deverá ser inicialmente avaliada.

A previsão das audiências televisivas poderá depender da caracterização da dinâmica dos comportamentos de consumo televisivo. As análises realizadas com os valores das séries temporais de determinados indicadores apenas deverá permitir que melhores resultados sejam alcançados, considerando o curto prazo e incidindo nas faixas horárias e nos indicadores que apresentam uma maior estabilidade.

Uma campanha publicitária deverá conseguir divulgar uma marca e incentivar a aquisição dos seus produtos e serviços. Nesse sentido, a actividade de planeamento de *media* deverá permitir uma selecção dos *media*, suportes e número de inserções publicitárias que permitam maximizar a exposição da publicidade ao segmento-alvo em função do orçamento disponível. Porém, os seus resultados estão, também, fortemente condicionados pelo grau de repetição da sua exposição aos telespectadores.

Quando se pretende construir uma campanha que repita a sua mensagem ao público-alvo, será necessário seleccionar os suportes publicitários que dupliquem entre si, mas no sentido da complementaridade dos conteúdos e não derivados de uma exposição meramente pontual e atribuída ao acaso. Adicionalmente, devemos conhecer a audiência acumulada, o que apenas será viável recorrendo à audimetria, uma vez que os anúncios podem ser repetidos algumas vezes no mesmo suporte. Isto é, os modelos de previsão devem tentar prever, não só a audiência de um programa e as correspondentes audiências comerciais, mas, também, o grau de duplicação das suas audiências.

A duplicação das audiências está fortemente condicionada pela repercussão das variáveis estruturais. A estruturação da grelha de programas condiciona a exposição televisiva e origina determinados comportamentos, como, por exemplo, a lealdade nas suas diversas vertentes, que podem, também, segundo alguns autores (*Ccooper, 1997*), contribuir para a previsão de quais serão as selecções realizadas no curto prazo. As variáveis estruturais impedem que as previsões sejam realizadas em função das preferências e dos conteúdos transmitidos e contribuem para que se produzam determinados comportamentos, sendo necessário avaliar a sua existência, interacção e a forma como se repercutem na evolução dos indicadores que definem o custo da publicidade televisiva.

A análise dos *ratings* é determinante para o planeamento de *media*, uma vez que contribui, igualmente, para a construção de modelos explicativos estruturais e, também, integrados da exposição televisiva, os quais, por sua vez, permitem avaliar a viabilidade e a identificação das melhores perspectivas de previsão dos dados. O desenvolvimento de modelos explicativos dos factores que influenciam o consumo televisivo proporciona informação que pode contribuir para a modificação da exposição ao *media* e para uma maior previsibilidade dos dados, assim como para a identificação de variáveis de previsão. No decorrer da construção das campanhas publicitárias, devemos procurar perceber como se produzem os fluxos de audiências entre os suportes publicitários seleccionados para integrar um plano.

As bases de audimetria permitem, também, a identificação de segmentos de telespectadores com perfis de consumo televisivo mais semelhantes e a sua posterior caracterização.

No planeamento de uma campanha publicitária, são necessárias três tipologias de modelos de *media*: os modelos de exposição, os modelos de optimização e os modelos de avaliação da repercussão da exposição aos suportes publicitários nos comportamentos dos consumidores e no volume de vendas alcançado. Na construção dos planos, deverá existir um “trade-off” entre os níveis de cobertura e de frequência, não existindo qualquer consenso relativamente aos níveis de frequência óptimos que devem ser adoptados.

O valor de um *media* para um anunciante é função dos objectivos tangíveis definidos no plano de *media*, por exemplo, se o canal consegue uma elevada cobertura do segmento-alvo e se os conteúdos programáticos são compatíveis com a imagem e posicionamento das suas marcas. As estruturas de preços da publicidade são função de factores tangíveis associados à dimensão das audiências históricas e futuras, da tipologia de programas que são transmitidos, dos segmentos sociodemográficos que estes conseguem cativar e da imagem global do canal (McMellon & Eftimov, 2004).

A construção de uma campanha publicitária não seria viável sem a previsão das audiências e a avaliação dos factores que condicionam a evolução de determinados indicadores que quantificam e caracterizam a exposição aos suportes, a nível programático e comercial. Em geral, as necessidades de informação são função do mercado em que se pretende adquirir o tempo publicitário. A informação mais relevante corresponde a se determinado conjunto de indivíduos contacta com um conjunto de suportes publicitários e com que frequência se verifica este contacto no decorrer de determinado período, independentemente de ser viável uma quantificação mais precisa da sua dimensão e dos indicadores utilizados nas análises.

Quer dizer, o termo “contactar” pode ser quantificado de diversas formas. Por vezes, é necessário reconstruir o comportamento de contacto dos indivíduos com o meio televisivo no decorrer de um horizonte temporal, o que facilmente poderá ser realizado recorrendo à audimetria. A partir desses dados, poderá ser necessário, numa segunda etapa, tentar prever esses comportamentos para o curto prazo ou de determinados indicadores.

Independentemente do nível de agregação dos dados considerado, as análises de previsão nas Ciências Sociais e, em especial, na área da audimetria apresentam inúmeras dificuldades. No passado, tem-se simplificado as análises com base no pressuposto de que a exposição histórica será equivalente à exposição futura, evitando, assim, a construção de modelos de previsão ou assumindo que a média dos valores históricos e os modelos de extrapolação são suficientes, existindo estabilidade nos dados. As análises de previsão são realizadas previamente à aplicação dos algoritmos de optimização, quando estes últimos são necessários. Posteriormente, existem alguns sistemas de apoio à decisão que permitem optimizar a construção dos planos, através da identificação do melhor conjunto de suportes aos quais associar as inserções publicitárias. A solução final estará condicionada por determinados objectivos de cobertura do alvo e considerando, simultaneamente, restrições orçamentais e sociodemográficas, mas, também, ao nível das tipologias e conteúdos programáticos ou dos canais a serem seleccionados. Contudo, esta solução encontra-se dependente da previsão da exposição e existem sempre factores que não são considerados pelo sistema e que dependem da experiência e conhecimento que o planificador detém a respeito da dinâmica dos comportamentos de consumo televisivo e que condicionam, por exemplo, as restrições que são impostas aos algoritmos.

A venda do espaço comercial é, também, realizada em função dos valores de *Gross Rating Points* (GRP) mínimos a garantir para o conjunto de anúncios, que, por sua vez, resulta do somatório dos valores de *ratings* para cada inserção contratada ou do produto da frequência e da cobertura alcançada. Por conseguinte, existem outras perspectivas de análise que viabilizam a aquisição de publicidade televisiva.

O investimento publicitário pode ser realizado em diferentes mercados, mas apenas na primeira fase, designada por “up-front market”, onde é vendida a maior parte das inserções do horário nobre, existindo o compromisso de garantir determinada exposição ao plano delineado, uma vez que, nesta fase, as receitas são mais elevadas. No “scatter market”, onde são vendidas as unidades remanescentes em planos, ou no “spot market”, onde é possível adquirir um menor conjunto de suportes um ou dois dias antes da sua transmissão a preços mais reduzidos, não são dadas garantias, por parte das estações televisivas, ao nível dos *ratings* que podem ser alcançados e quanto à sua composição sociodemográfica. Consequentemente, as agências de *media* devem dispor desta informação para garantir que sejam alcançados os melhores resultados.

A partir dos resultados das análises de previsão, seleccionam-se as combinações consideradas óptimas de blocos comerciais que serão mais rentáveis para um anunciante. Como passo prévio à aplicação de qualquer sistema de optimização do investimento, é necessário conhecer a audiência dos programas e blocos comerciais onde possa investir-se. As análises devem ser realizadas da forma mais rápida possível, a fim de que as unidades sejam adquiridas logo após a publicação das grelhas de programação, devendo ser privilegiadas as abordagens quantitativas de previsão. Os indicadores mais relevantes para o sector correspondem aos que quantificam a duplicação de audiências entre os suportes e aos valores de *ratings*, que, por sua vez, dependem dos valores de *shares* de audiências e da exposição global ao *media*, quantificada através de um indicador designado por taxas médias de audiências para o total da TV.

Portanto, é necessário avaliar a viabilidade de previsão dos principais indicadores por aplicação de diferentes abordagens e promover a identificação de outras perspectivas de análise que possam contribuir para a previsão das audiências em função dos diversos contextos que podem ser identificados na evolução dos dados. Os contextos resultam das necessidades de informação das estações televisivas ou das agências de meios, e diferenciam-se quanto à evolução e variabilidade dos dados e ao nível das metodologias e recursos envolvidos.

Os modelos de previsão podem ser qualitativos ou quantitativos e, neste último contexto, univariados ou multivariados, lineares ou não lineares, explicativos e causais ou univariados; podem ser modelos de Regressão, econométricos ou de *Data Mining* (DM), entre outras classificações. Existem inúmeros contextos que condicionam as abordagens a adoptar. As previsões quantitativas são essenciais para o planeamento de *media*, enquanto as estações televisivas podem recorrer a previsões qualitativas quando se pretende quantificar a audiência de um novo programa ou tipologia ou quando se altera a sua posição na grelha. Embora os erros estejam sempre presentes nas previsões quantitativas, esta perspectiva de análise permite planear de forma mais precisa, quando comparada com o planeamento baseado unicamente na experiência e na intuição de determinados profissionais.

Os modelos quantitativos “causais” apresentam a capacidade de poderem efectuar previsões a médio prazo, considerando variáveis exógenas. Quando as variáveis explicativas não são conhecidas ou previsíveis, estes modelos devem ser designados por modelos explicativos, em vez de modelos de previsão. Um modelo de Regressão assume que as variáveis explicativas são deterministas e conhecidas, mas, no âmbito da audimetria, são inúmeras as variáveis que não se inserem nesta classificação e que influenciam a exposição e os indicadores que dela resultam. Em Ciências Sociais, dificilmente conseguem ser definidas relações de causa e efeito. Consequentemente, é necessário avaliar a viabilidade dos modelos de *ratings* e identificar perspectivas alternativas de análise.

Os *ratings* abrangem, em sentido lato, a totalidade dos dados que podem ser obtidos através do processo de audimetria e, a um nível mais específico, um indicador que quantifica a percentagem do público útil que se encontra exposto a determinado suporte. Os modelos de *ratings*, ao contrário dos “viewing choice models”, caracterizam a exposição televisiva do ponto de vista agregado, ou seja, incidem nos fluxos de audiências entre os suportes. Esta vertente de análise associa-se à teoria estrutural de análise da exposição televisiva, à investigação quantitativa, a um macro nível de análise e aos instrumentos de registo da exposição, destacando-se a audimetria.

Outras perspectivas, nomeadamente a denominada de gratificações e usos do *media*, no âmbito da Psicologia e da Sociologia, incidem no processo de decisão individual e nos motivos que para tal contribuem, entre outras vertentes de análise. Por oposição, os modelos de *ratings* procuram caracterizar e compreender o modo como os factores estruturais e as variáveis sociodemográficas, que constam das bases de audimetria, influenciam a evolução de indicadores programáticos e comerciais que são essenciais para o sector. Simultaneamente, estes modelos podem avaliar a viabilidade da sua previsão ou contribuir para a definição das abordagens que podem conduzir aos melhores resultados a este nível.

I.3. A QUESTÃO DE PARTIDA

O problema da investigação pode ser centrado, a nível global, na questão que a seguir se apresenta, e que reflecte a avaliação do contributo de áreas metodológicas mais adequadas e recentes, face ao que consta da bibliografia da especialidade, para a construção dos modelos que condicionam a aquisição de publicidade televisiva. A actividade de planeamento de *media* não pode ser realizada sem recorrer às bases de audimetria ou a um sistema de informação equivalente ainda que, eventualmente, possa ser complementada com fontes de dados alternativas em determinadas etapas deste processo.

➤ Qual o contributo das metodologias de *Data Mining* e de *Análise Simbólica* para o desenvolvimento de modelos de *ratings*, enquanto tipologia de dados associada ao processo de audimetria, que contribuam para a construção dos modelos de exposição que são essenciais para a tomada de decisão a nível operacional e estratégico no sector televisivo?

O planeamento de *media* pode ser realizado por previsão de valores de *ratings*, mas, também, por quantificação dos fluxos agregados de audiências entre os suportes publicitários seleccionados para integrar determinada campanha publicitária. Pretende-se a identificação das perspectivas de análise que podem conduzir aos melhores resultados no âmbito dos modelos que visam a previsão das audiências televisivas para o curto prazo.

Consequentemente, é necessário conhecer, previamente, a evolução dos indicadores agregados e acumulados de audimetria. O objectivo geral pode ser desagregado na avaliação de quais são os factores de que dependem maioritariamente os *ratings*, o que implica a necessidade de caracterizar a evolução dos *ratings* programáticos e comerciais, ou seja, os indicadores que condicionam, maioritariamente, a construção das grelhas de programas e dos planos de *media* e a venda do espaço comercial. É, igualmente, necessário proceder à análise dos “factores de bloco” e à caracterização de determinados “comportamentos estruturais” de consumo televisivo, quantificados através dos indicadores acumulados de audimetria.

No presente estudo, avalia-se a viabilidade de previsão das audiências televisivas nos contextos que condicionam, maioritariamente, a venda do espaço comercial e considerando metodologias mais adequadas às características que apresentam as bases de dados de audimetria. As perspectivas mais adequadas são função da evolução dos indicadores e do grau de regularidade que caracteriza o consumo televisivo. Em Ciências Sociais, os modelos não têm como função prever os comportamentos e variáveis de forma determinista, mas, sim, explicar e, eventualmente, conseguir prever o máximo que for possível das variáveis observadas.

As análises decorrentes deste trabalho devem contribuir para a identificação das abordagens que podem conduzir aos melhores resultados, ainda que o seu grau de precisão não possa ser equivalente ao que é exequível nas Ciências Exactas. Quando consideramos os dados agregados, ou seja, os *ratings*, que resultam da exposição de um painel de audimetria, deve ser adoptada a mesma perspectiva.

Quer dizer, as análises devem identificar os factores controláveis que condicionam a evolução dos indicadores e as influências mais relevantes sobre as quais uma organização poderá ou não ter poder de decisão. Devemos avaliar se as correlações observadas correspondem a reais relações de causa-efeito, a associações ou a meras coincidências observadas nos dados e se as relações identificadas são compreensíveis e se estão de acordo com o senso comum.

Com base nos resultados, devemos identificar quais são as melhorias que podem ser alcançadas no desempenho organizacional e qual o conhecimento que pode ser retirado da análise dos erros de previsão. As análises de previsão nas Ciências Sociais e nas Ciências Exactas apresentam características que as diferenciam significativamente.

O objectivo das análises realizadas no âmbito deste trabalho reside na avaliação das influências estruturais na evolução dos dados que condicionam a construção de uma campanha publicitária e na viabilidade da sua previsão, identificando as perspectivas mais adequadas, que são função dos recursos disponíveis e da evolução dos dados.

O investimento em publicidade televisiva poderá ser realizado por previsão de valores das taxas médias de audiências programáticas e comerciais, mas não só. Atendendo às grandes dificuldades que as análises de previsão colocam neste domínio, pretendemos avaliar a capacidade predictiva de determinadas perspectivas explicativas da exposição e a construção de um modelo mais abrangente face aos que constam da bibliografia da especialidade. Pretendemos demonstrar os múltiplos factores que condicionam o consumo televisivo e as suas implicações ao nível da construção dos modelos de exposição. Devemos, também, avaliar a repercussão dos indicadores acumulados na evolução dos *ratings*.

Quando as análises explicativas identificam factores conhecidos *a priori* ou previsíveis, estes podem ser utilizados, posteriormente, em modelos de previsão do indicador em análise. A análise dos “comportamentos estruturais” pretende caracterizar a exposição televisiva segundo determinadas dimensões que condicionam a venda do espaço comercial. Esta análise apresenta, como objectivo final, avaliar se a audiência é variável e pouco consistente quando se relacionam os indivíduos e os suportes, uma vez que a dinâmica dos comportamentos condiciona a viabilidade da previsão da exposição e a componente metodológica a adoptar em determinados contextos.

Os planos de *media* são construídos considerando a cobertura e a frequência média ou efectiva, entre outras classificações, não existindo, de acordo com a bibliografia da especialidade, consenso relativamente aos valores “óptimos” das frequências de exposição. Recorrendo às análises de duplicação das audiências, avaliam-se os suportes a incluir nos planos, uma vez que a maioria das inserções comerciais é transaccionada de forma agregada.

Neste contexto, é relevante avaliar de que dependem, maioritariamente, as audiências acumuladas, considerando abordagens analíticas que permitam, também, captar a não linearidade das relações e recorrendo a fontes de dados mais adequadas à sua análise. Em função dos resultados, avaliar-se-á a capacidade predictiva destes comportamentos e as abordagens analíticas que possam contribuir, neste contexto, para a previsão da exposição.

As análises de previsão devem ser realizadas de modo a privilegiar-se uma maior rapidez e um menor investimento financeiro. Consequentemente, a previsão dos *ratings* constitui uma das principais perspectivas de análise, sendo necessário avaliar quais são as abordagens metodológicas que permitem alcançar os maiores ganhos ao nível da precisão dos resultados. Esta perspectiva deverá, também, ser privilegiada, uma vez que corresponde à única tipologia de dados a que tem acesso a maioria das agências de meios.

A *Figura I.3.1* ilustra a sequência de análises que viabiliza a avaliação das melhores abordagens de construção dos modelos de exposição.

Em função dos resultados, através dos quais se constatará qual é o grau de estabilidade dos indicadores e a regularidade da exposição agregada, será possível concluir se a previsão dos *ratings* constitui uma opção a considerar, e em que contextos, e apresentar alternativas de análise que podem ser função dos modelos explicativos dos comportamentos de consumo televisivo do painel de audimetria.

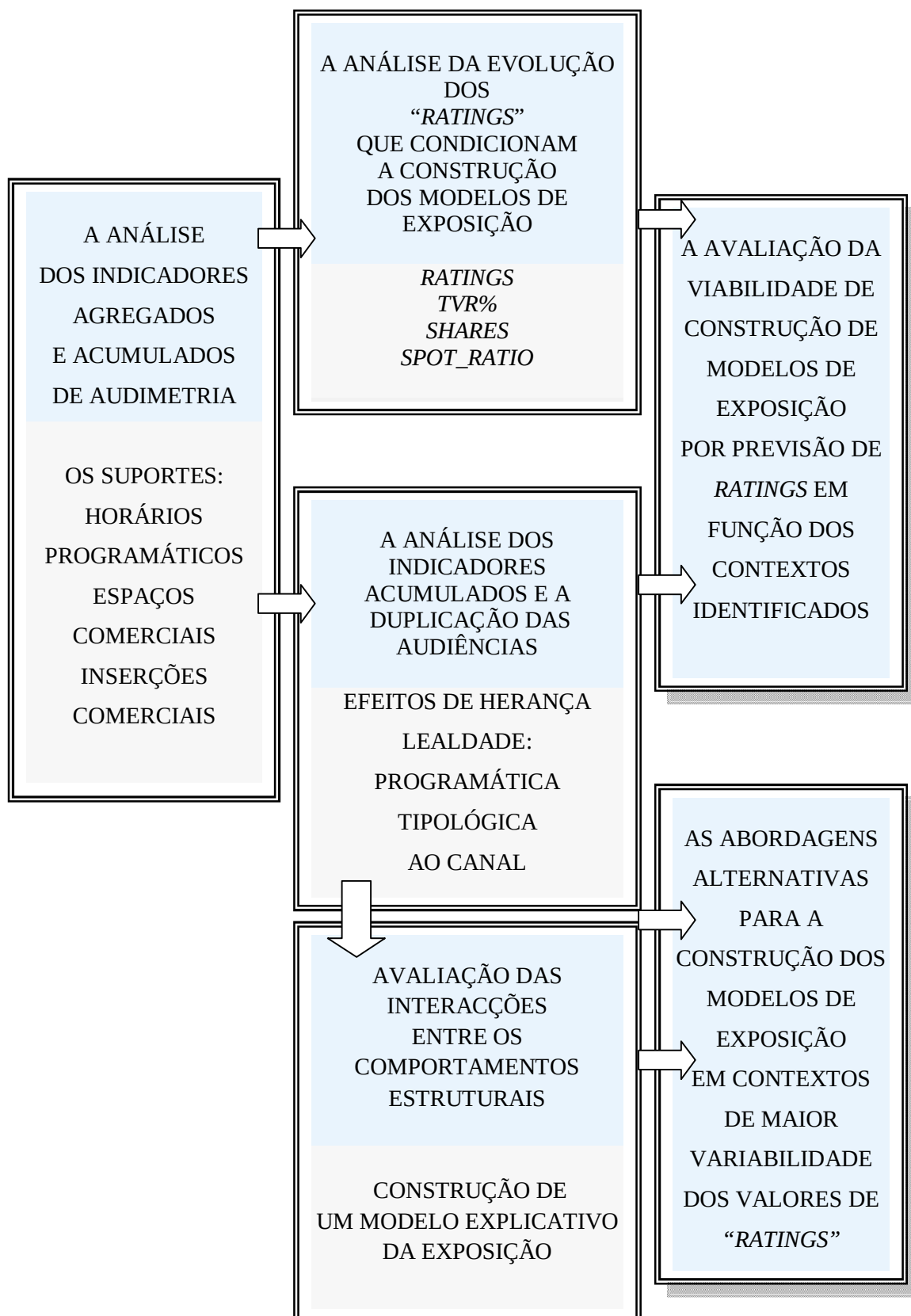


Fig. 1.3.1 - Sequência de análises que viabilizam a apresentação de uma proposta para a previsão das audiências televisivas. Fonte: Elaborado pela autora.

I.4. AS PRINCIPAIS CONTRIBUIÇÕES DA TESE

As análises a realizar pretendem contribuir para as inúmeras lacunas existentes na área das análises explicativas e de previsão da exposição televisiva e dos principais indicadores de audimetria, quer ao nível dos métodos analíticos adoptados, quer das questões a desenvolver. A nível nacional, apenas foi viável identificar um reduzido número de trabalhos na área dos modelos de *ratings*, como, por exemplo, a previsão de valores de *shares* de audiências por aplicação de Árvores CART (*Datia, 2002*) e realizadas apenas com dados de audimetria do período 1996/1997, onde se verifica uma maior estabilidade na sua evolução. Em *Fernandes (2001)*, consta uma exposição sobre a informação que pode ser extraída das bases de dados de audimetria e as referências bibliográficas apresentadas são apenas internacionais. Por conseguinte, atendendo à dificuldade de identificação de publicações que incidam na análise das bases de dados nacionais, as incógnitas e dificuldades são, inevitavelmente, superiores.

Na presente tese consta um número substancial de análises exploratórias e descritivas da evolução dos índices de audiência programáticos e comerciais, assim como da exposição televisiva, nomeadamente no âmbito da duplicação das audiências, em que se avalia o grau de regularidade da exposição ao *media* e aos suportes, os factores que o condicionam e as respectivas interdependências. A duplicação das audiências inclui a repetição de selecções, os efeitos de herança e a lealdade ao canal e tipológica, permitindo avaliar a viabilidade da previsão da exposição individual e agregada.

Perante a dificuldade de identificação de publicações recorrendo às bases de dados nacionais, foi necessário realizar um abrangente conjunto de análises exploratórias de caracterização da evolução dos principais indicadores de audimetria e de alguns comportamentos de exposição ao *media*, por forma a viabilizar a apresentação de uma proposta que possa contribuir para a temática da previsão das audiências, para o planeamento de *media* e para uma maior compreensão de como os telespectadores interagem com o televisor.

As previsões nunca serão exactas, uma vez que resultam de comportamentos humanos, os quais são classificados como imprevisíveis, existindo uma permanente influência da envolvente externa e social do indivíduo e de estímulos internos, frequentemente, ambíguos e não quantificáveis nem previsíveis. Apenas podemos apresentar e comparar as potencialidades de possíveis abordagens quantitativas e/ou qualitativas de previsão das audiências de programas e comerciais, que serão função dos contextos em análise, contribuindo, assim, para a optimização do investimento publicitário e para a gestão das empresas ligadas a este sector de actividade.

Pretende-se avaliar em que medida a *Análise Simbólica* e o *Data Mining* podem contribuir para que sejam alcançados melhores resultados no âmbito dos modelos de *ratings*. Uma vez que as conclusões de trabalhos internacionais no âmbito estrutural não se ajustam à realidade nacional, realizou-se um conjunto considerável de análises estatísticas recorrendo a abordagens metodológicas inexistentes ou muito pouco exploradas na bibliografia internacional. Foram realizadas análises do âmbito da previsão, Regressão, classificação, factoriais e agrupamento de dados recorrendo à Estatística Multivariada, à *Análise Simbólica* e ao *Data Mining* Temporal, Proposicional e Sequencial que, atendendo aos dados que foram disponibilizados, permitam contribuir para a análise dos principais indicadores de audimetria.

Segundo a bibliografia da especialidade, as análises denominadas de explicativas diferenciam-se das análises de previsão por incidirem na avaliação da repercussão de factores estruturais nos índices de audiência, mas que nem sempre são previsíveis (*Weber, 2002*). O consumo televisivo encontra-se influenciado por inúmeros factores e por um conjunto de variáveis que se inter-relacionam, complementando-se, compensando-se e intensificando os respectivos efeitos de forma variável no decorrer do tempo. Consequentemente, dificilmente será viável a identificação de factores e interdependências que se mantenham válidos por um grande horizonte temporal, devendo ser reavaliados sempre que existam significativas alterações nas condições subjacentes.

As variáveis explicativas serão relevantes do ponto de vista da previsão se os seus valores forem conhecidos ou calculados com alguma fiabilidade no momento de realizar uma previsão. Os modelos explicativos de *ratings* baseiam-se, essencialmente, na avaliação dos factores estruturais, abrangendo, também, a análise de comportamentos de duplicação, objecto de estudo neste trabalho, por aplicação de metodologias que não constam da bibliografia da especialidade, nomeadamente as abordagens não lineares e por avaliação da repercussão de um conjunto muito mais abrangente de variáveis que foram calculadas a partir da informação que consta das bases de audimetria.

As análises que visam uma maior percepção sobre a forma como as variáveis estruturais influenciam a exposição televisiva são determinantes para uma futura previsão dos índices de audiências. Previamente ao desenvolvimento de um modelo de previsão, será necessária a caracterização da evolução das variáveis em estudo, a identificação dos factores que condicionam a exposição da audiência e a avaliação das respectivas interdependências. A previsão corresponde à temática que apresenta uma maior complexidade em Ciências Sociais; logo, os trabalhos encontram-se direccionados, em geral, apenas para a explicação dos fenómenos observados. Atendendo ao conjunto muito diversificado e abrangente de análises que são necessárias para procurar caracterizar e explicar os padrões de consumo televisivo e viabilizar a construção dos planos de *media*, para as análises que não foi viável realizar, por falta de recursos e atendendo às condicionantes temporais existentes, mas que se apresentam como relevantes, constam as abordagens metodológicas mais apropriadas, uma vez que, sem métodos analíticos, nenhuma informação relevante poderá ser extraída de qualquer base de dados.

A presente tese procurou promover a integração de metodologias de diversos âmbitos da análise estatística de dados, assim como da teoria da comunicação social, por forma a contribuir para a temática dos modelos de exposição, ainda que privilegiando a componente analítica. A previsão das audiências televisivas encontra-se fortemente dependente das tecnologias de medição das audiências, das abordagens metodológicas e de manuseamento de grandes bases de dados e dos modelos explicativos da exposição.

Relativamente às fontes de dados, um painel corresponde à metodologia idónea para poder medir-se a evolução dos comportamentos e dos indicadores que deles resultam. O que é declarado não cumpre os requisitos do que é medir, onde deverá existir objectividade. Será, inevitavelmente, o desenvolvimento dos sistemas de informação que irá, certamente, permitir que melhores resultados sejam alcançados no futuro, sendo esta a fonte de dados utilizada para a realização das análises propostas neste trabalho. A audimetria viabiliza a construção de inúmeros indicadores que não seriam viáveis recorrendo aos métodos alternativos de medição das audiências.

O *Data Mining* define o processo automatizado de análise de grandes bases de dados para extrair informação de apoio à gestão das organizações, sendo as metodologias utilizadas, quer para descrever características passadas, quer para prever tendências e regularidades que se pressupõe manterem-se válidas no futuro. O *Data Mining* foi uma das áreas privilegiadas, procurando avaliar-se o seu contributo para que melhores resultados sejam alcançados nas diversas etapas que viabilizam a construção de uma campanha publicitária.

As análises de duplicação das audiências realizadas pretenderam avaliar se existem as regularidades que são essenciais ao *Data Mining* e que condicionam a viabilidade de serem realizadas análises de previsão. De acordo com *Hand (1998)*, a Estatística preocupa-se com a análise primária dos dados, no sentido de que as bases de dados são construídas de acordo com determinado objectivo ou para permitir a avaliação de um conjunto de questões que são colocadas *a priori* pelo investigador. Em *Branco (2010)* é referido que o *Data Mining*, por outro lado, incide, também, numa análise “secundária” dos dados, isto é, persegue uma perspectiva mais abrangente e indutiva face ao que se verifica numa abordagem hipotético-dedutiva, que tem sido, frequentemente, considerada como a base do progresso da ciência moderna. Assim, o *Data Mining* pode ser visto como um descendente directo da Estatística, já que se trata de um conjunto de técnicas metodológicas que, na realidade, devem ser utilizadas de forma complementar. Porém, as cooperações entre estes dois âmbitos são reduzidas, mas devem coexistir no campo da audimetria, para que melhores resultados possam ser alcançados.

A análise das audiências televisivas beneficiará, certamente, com a aplicação de ambas as áreas, apesar das grandes dificuldades que as análises de indução colocam às Ciências Sociais, inviabilizando que resultados satisfatórios possam ser alcançados, ao contrário do que se verifica, frequentemente, nas Ciências Exactas, onde as previsões com um carácter “determinista” podem ser viáveis. O determinismo estrito defende que o estado presente do universo ou que os estados do passado e do presente fixam, de maneira unívoca, o estado do universo em qualquer instante do futuro. Porém, as previsões nestes contextos podem, também, não ser viáveis.

Segundo a Mecânica Clássica, o determinismo estrito é, igualmente, válido para qualquer sistema completamente isolado do resto do universo, quer dizer, para um sistema cuja evolução não é afectada de maneira significativa pelo ambiente, portanto, por circunstâncias e condições que, certamente, não se verificam nesta área, onde comportamentos de pessoas estão em análise e continuamente sujeitos a uma diversidade de estímulos externos e internos, e variáveis no decorrer do tempo, o que coloca grandes incógnitas relativamente à viabilidade da previsão dos indicadores de audimetria. Verificam-se, frequentemente, limitações e dificuldades equivalentes em outros domínios, como, por exemplo, na área financeira. No entanto, o determinismo, a causalidade e a previsibilidade não são, também, conceitos que possam ser considerados como equivalentes.

Por conseguinte, foi necessário avaliar qual é grau de regularidade da exposição televisiva e as suas implicações ao nível da previsibilidade da exposição agregada e dos índices de audiências. Assim, recorrendo às análises de duplicação das audiências, pretendeu avaliar-se quais são as perspectivas de análise que podem conduzir aos melhores resultados possíveis, e não a previsões exactas, que não são claramente viáveis neste domínio. Atendendo às análises realizadas, é evidente a complexidade existente ao nível do desenvolvimento das análises explicativas e de previsão da exposição televisiva e dos índices de audiências programáticos e comerciais, assim como em relação à segmentação do painel de audimetria.

Quer as metodologias de *Data Mining*, quer de Estatística Multivariada, incluindo os equivalentes algoritmos simbólicos que permitem a agregação dos dados, podem contribuir positivamente para o desenvolvimento de modelos de *ratings*, apesar de os resultados publicados em função dos contextos nem sempre serem satisfatórios, atendendo às necessidades de informação do sector televisivo. Não obstante a separação que se observa no campo metodológico de análise de dados, a presente tese descreve e demonstra alguns potenciais de aplicação de ambas as áreas, uma vez que, frequentemente, se defende a inadequabilidade das mesmas para a análise de bases de dados de audimetria. Certamente que existem outras metodologias que podem contribuir para esta área no âmbito da Investigação Operacional ou da Engenharia Informática, entre outras áreas.

Sendo o objectivo das análises, por vezes, a previsão, mas, em maior grau, a explicação da evolução dos índices de audiências, torna-se inevitável recorrer à audimetria. Esta poderá ser, posteriormente, complementada com outras fontes, em função dos contextos, os quais se descrevem, também, na presente tese. No entanto, são as teorias estruturais do consumo televisivo associadas à investigação mais quantitativa e aos dados de *ratings* que têm viabilizado os melhores resultados a nível internacional. Os métodos qualitativos produzem maiores erros e as restantes perspectivas explicativas do comportamento dos telespectadores e fontes de dados são insuficientes, atendendo às necessidades de informação que o planeamento de *media* requer.

As sondagens de intenção de ver não têm apresentado resultados satisfatórios e as restantes abordagens qualitativas apresentam mais-valias quando se pretende a construção das grelhas, mas não para o planeamento de *media*. A dificuldade inerente à previsão de determinados indicadores não resulta das limitações desta tecnologia, mas, sim, da própria “imprevisibilidade” da vida real e dos comportamentos humanos, bem como da forte variabilidade que podem apresentar os indicadores que deles resultam. Uma vez que um grande número de telespectadores oculta a realidade ou substitui uns programas por outros, os inquéritos acabam por não ter grande valor para este sector, para além de não constituírem um instrumento de previsão.

Poderá recorrer-se, também, a um painel de opinião, que, em função da sua intuição e experiência, sugere possíveis margens de valores para os indicadores que condicionam a venda do espaço comercial. As vantagens e as desvantagens das diversas perspectivas foram descritas nesta tese, ficando evidente a necessidade de os modelos quantitativos serem desenvolvidos através da audimetria.

Os resultados permitiram-nos constatar que uma menor percentagem do painel poderá apresentar uma exposição mais regular ao meio televisivo. Existindo alguma regularidade e determinados hábitos de consumo televisivo, a previsão será viável, com maiores ou menores margens de erro, desde que recorrendo à audimetria e às adequadas abordagens metodológicas que dependem dos contextos e que serão apresentadas e descritas ao longo desta tese. Quando se verifica uma maior irregularidade na exposição, dificilmente será viável conseguir prever a disponibilidade dos telespectadores e, posteriormente, a afectação do seu tempo livre ou tempo vazio ao *media* e aos diversos suportes disponíveis em determinado momento. Uma vez que o comportamento individual é imprevisível, foram avaliadas perspectivas alternativas de análise para a construção dos planos de *media*.

Através da audimetria, para quem apresenta uma exposição televisiva mais irregular, não será viável saber se num dia e a uma hora específica determinado telespectador irá ligar o televisor. Não existindo uma fidelização para com o meio televisivo, os comportamentos estruturais devem apresentar uma menor influência. A disponibilidade e o desejo de ver televisão podem ser os principais factores responsáveis pela variabilidade dos dados. Posteriormente, as selecções podem corresponder aos programas que os telespectadores considerem mais apelativos ou menos entediantes no momento em que decidem ligar a TV. Se a audimetria não permite responder a estas questões, os estudos de memória e os estudos coincidentais também não, e interrogar os telespectadores a respeito das suas selecções futuras não será, igualmente, uma melhor opção. Por isso mesmo, em especial a previsão das audiências de eventos desportivos não tem sido viável, uma vez que depende da capacidade de determinado evento desportivo para conseguir atrair telespectadores que, em geral, não

ligam o televisor nessas faixas horárias e, posteriormente, conseguir prever a sua repercussão na programação adjacente. Conseguir prever comportamentos concretos no espaço e no tempo para os segmentos que assistem, pontualmente, e de forma muito irregular, ao *media*, será dificilmente viável, com ou sem audimetria. Porém, atendendo à dimensão deste segmento, poderá não ser viável a construção de uma campanha publicitária sem ter, pelo menos parcialmente, em consideração estes telespectadores. Se os indicadores de audimetria apresentarem alguma estabilidade, esta será a melhor perspectiva de análise para promover a construção dos planos de *media*; porém, estes apresentam, também, com frequência, uma evolução que o inviabiliza. Existindo estabilidade ao nível dos *ratings*, a previsão univariada ou multivariada de séries temporais e os modelos de Regressão linear e não linear devem ser testados e avaliadas as melhores perspectivas de análise, apresentando-se algumas sugestões de análise na presente tese.

Por conseguinte, foi, também, descrita uma proposta na área do *Data Mining* não proposicional que permita ultrapassar, em parte, as dificuldades que estas análises colocam ao sector televisivo quando a exposição e os indicadores que dela resultam são fortemente irregulares. Não se pretende promover previsões com uma elevada precisão, mas, sim, contribuir para a selecção dos suportes mais adequados aos objectivos de uma campanha publicitária e minimizar os custos associados ao diferencial observado entre os valores previstos e os valores reais.

Se considerarmos a fracção do painel que assiste com alguma regularidade ao meio televisivo, recorrendo aos dados de audimetria será mais fácil tentar prever as suas futuras selecções com base no seu histórico de exposição televisiva, que a audimetria descreve com um maior rigor face aos métodos alternativos existentes. Para este segmento de telespectadores, será relevante verificar a forma como as variáveis e os comportamentos estruturais influenciam o consumo televisivo. Será necessário proceder à avaliação dos comportamentos de lealdade e de efeitos de herança, entre outras perspectivas de análise que são abordadas nesta tese.

Na área da análise dos efeitos de herança e da lealdade, o recurso a inquéritos por entrevista e por questionário não permite determinar quantos telespectadores transitam de programa para programa e durante quanto tempo vão assistir a cada um dos programas envolvidos. De igual forma, não será viável quantificar quantos telespectadores vão, no decorrer de determinado programa, transitar para um canal concorrente e quais os instantes em que isso se verifica ou quais os instantes em que vão ligar e desligar o televisor. Estes comportamentos estruturais apenas podem ser analisados com um maior rigor quando são utilizados os dados de audimetria. Recorrendo a esta tecnologia, podemos, através das adequadas abordagens metodológicas, tentar verificar se os telespectadores apresentam alguma tendência para permanecer no mesmo canal de programa para programa e se as suas selecções programáticas incidem em maior grau em determinados programas, tipologias ou canais. Consequentemente, tentar extrair o máximo de informação das bases de dados de audimetria será a melhor perspectiva para tentar prever e perceber a evolução de determinados indicadores para alguns segmentos, mas não para a sua totalidade. Ao nível dos dados dos ficheiros de base de audimetria, estes acrescentam, também, grandes mais-valias se o objectivo das análises for a caracterização da dinâmica subjacente às selecções programáticas e à exposição ao *media*.

A previsão de determinados indicadores nem sempre implica, obrigatoriamente, a necessidade de uma caracterização mais detalhada do processo de consumo televisivo, ou seja, o desenvolvimento dos denominados modelos explicativos de audiências. Isto verifica-se, essencialmente, nos contextos de maior variabilidade dos dados. Para o desenvolvimento de modelos explicativos de audiências, a análise dos dados de audimetria apresenta, também, grandes mais-valias, podendo, posteriormente, se necessário, ser complementada com outras abordagens de carácter mais qualitativo. Em função das sequências programáticas associadas a cada telespectador, podem ser retiradas ilações quanto às suas motivações; por exemplo, se sistematicamente alguém apresenta um padrão de exposição que consiste meramente na selecção de espaços informativos, podemos concluir que a vertente informativa será a privilegiada e que existe um reduzido envolvimento com o meio televisivo.

No presente trabalho, do ponto de vista explicativo, são abordados os diversos comportamentos estruturais de consumo televisivo. Para a realização de análises sociológicas, que pretendam promover uma melhor percepção a respeito de outros aspectos relacionados com o consumo televisivo, os dados de audimetria não serão, obviamente, os únicos que devem ser utilizados, mas podem e devem, certamente, ser vistos como complementares, atendendo às mais-valias que acrescentam às análises.

Os sistemas de medição das audiências registam comportamentos que podem ser aprofundados, quando necessário, por parte de outras disciplinas e abordagens. No entanto, não se pretendeu, também, avaliar determinadas questões de cariz mais qualitativo. A necessidade de introduzir a temática dos modelos explicativos da exposição resulta do facto de estes modelos serem essenciais, em determinados contextos, para as análises de previsão necessárias à construção dos planos de *media*.

A nível prático, as análises de previsão podem colocar inúmeras dificuldades, como, por exemplo, não se dispor, por vezes, da listagem dos programas que cada canal vai transmitir com a necessária antecedência para viabilizar a realização dos cálculos em tempo útil. Por outro lado, os canais alteram, também, com frequência, a muito curto prazo a constituição da grelha que foi anteriormente divulgada. A excessiva concorrência entre os canais poderá tornar difícil o acesso aos dados que descrevam com alguma fiabilidade os programas a transmitir no futuro.

Todavia, a nível nacional, nos períodos de maiores audiências, algumas destas dificuldades nem sempre se verificam, uma vez que os canais recorrem à transmissão de séries e telenovelas no decorrer de um longo horizonte temporal. Além do mais, quanto mais os telespectadores apresentarem uma exposição estrutural, menor será a repercussão da variabilidade dos conteúdos e mais previsível será o consumo televisivo, desde que exista alguma regularidade na exposição ao *media*. Consequentemente, existe uma forte interdependência entre os modelos explicativos e as análises de previsão das audiências.

Foi, também, proposta e descrita uma abordagem a ser utilizada na análise das audiências dos jogos de futebol ou quando novos conteúdos são introduzidos na grelha, sendo, neste último contexto, maioritariamente função do perfil de exposição do segmento-alvo e de imperar uma selecção mais intencional ou menos intencional.

Quando se verifica uma maior interacção entre as séries temporais multivariadas dos diversos indicadores, como, por exemplo, nos contextos em que são transmitidos programas desportivos, devem ser avaliadas perspectivas de análise que permitam compreender a evolução dos dados e deve tentar-se prever a sua direcção futura mais provável, para que os anunciantes ou os seus intermediários possam realizar uma melhor negociação com as estações televisivas. Porém, frequentemente, pelas limitações de orçamento que os anunciantes apresentam, os planificadores optam por suportes menos dispendiosos, ainda que estes apresentem uma menor cobertura do mercado-alvo.

I.5. A ESTRUTURA DA TESE

Atendendo à grande diversidade de temáticas que imperam no sector televisivo e à falta de articulação do conjunto, procurou fazer-se um levantamento dos aspectos teóricos e práticos que envolvem o comportamento dos telespectadores.

Foram considerados os contextos que são mais relevantes para o planeamento de *media*, assim como a respeito da sua previsibilidade ou dos índices de audiências que deles resultam. Apresenta-se uma introdução de enquadramento das análises propostas, bem como à actividade de planeamento de *media*, aos modelos explicativos parciais e integrativos da exposição ao TV e às análises de previsão das audiências televisivas considerando os índices de audiências, o comportamento agregado e o individual.

O estado da arte incide nas diversas perspectivas de análise que condicionam a construção dos modelos de exposição, ou seja, nas análises explicativas e/ou de previsão dos principais índices de audiências e dos comportamentos “estruturais” que condicionam a selecção dos suportes e as metodologias de previsão mais adequadas ao planeamento de *media*.

A revisão da literatura permitiu identificar, essencialmente, referências teórico-empíricas internacionais, sendo muito escassas as publicações nacionais ao nível da análise das bases de audimetria e, em especial, no âmbito dos modelos de *ratings* e da análise dos comportamentos estruturais de consumo televisivo. Uma vez que existe uma forte influência dos contextos geográficos e temporais, as referências identificadas acabam por não contribuir para um maior conhecimento do que o que se verifica a nível nacional.

Inevitavelmente, a fonte de dados utilizada corresponde às bases de dados obtidas segundo o processo de audimetria que viabilizam uma análise diacrónica da audiência e a caracterização da exposição com uma maior objectividade, precisão e detalhe. Apresenta-se uma descrição e comparação dos diferentes métodos de medição das audiências, salientando as respectivas vantagens e limitações.

Do ponto de vista metodológico, recorre-se a abordagens que não constam ou que são escassas na bibliografia internacional, nomeadamente as análises não lineares associadas a diversas subáreas do *Data Mining* e à Análise Simbólica, que representa a extensão das metodologias multivariadas de análise de dados a bases de dados simbólicas, onde uma observação não corresponde apenas a um único valor quantitativo ou qualitativo. Recorrendo a estas áreas, pretende-se avaliar se melhores resultados podem ser alcançados do ponto de vista dos modelos de *ratings*. Em função dos resultados, constam propostas alternativas de análise que possam contribuir para a construção dos planos de *media* e para uma maior percepção sobre a forma como o painel interage com o televisor.

No capítulo de apresentação dos principais resultados constam algumas análises exploratórias, descritivas e de previsão dos principais indicadores de audimetria programáticos e comerciais, recorrendo a metodologias univariadas e multivariadas. Os indicadores em análise correspondem aos *ratings*, às taxas médias de audiências para o total da TV e aos *shares* de audiências. No âmbito da duplicação das audiências, considera-se a lealdade programática, tipológica e ao canal, assim como os efeitos de herança que podem ser analisados recorrendo aos indicadores acumulados.

Posteriormente, em função dos resultados, consta uma proposta que possa contribuir para as necessidades de informação que a construção dos modelos de exposição coloca ao sector. A estrutura da presente tese, representada na *Figura I.5.1*, preenche as lacunas existentes na bibliografia da especialidade e segue de perto as necessidades de informação inerentes ao desenvolvimento dos modelos de exposição.

➤ No *Capítulo I* faz-se a descrição do problema, definem-se as fronteiras do problema em estudo e o objectivo geral a desenvolver, caracteriza-se a estrutura da tese e salientam-se as suas principais contribuições.

➤ Os objectivos gerais e específicos, as motivações e as questões operacionais a desenvolver constam do *Capítulo II*.

➤ Uma descrição mais exaustiva das diferentes linhas de investigação existentes *no âmbito* da análise da exposição televisiva e a descrição de alguns conceitos associados ao sector televisivo e ao planeamento de *media* constam dos *Apêndice B a D* e, mais sinteticamente, no *Capítulo III*.

No *Apêndice B* caracteriza-se o sector televisivo nacional e as vantagens em recorrer à publicidade televisiva. Caracteriza-se a actividade de planeamento de *media* que permite a planificação e avaliação das campanhas publicitárias. Neste apêndice, consta a descrição das diversas etapas que integram a construção de um plano de *media* e dos actores intervenientes, assim como dos objectivos gerais do planeamento de *media*,

informação necessária e métodos de construção dos planos. Pretende-se identificar e justificar a necessidade do desenvolvimento de análises explicativas e de previsão da exposição televisiva, assim como dos indicadores de audimetria e descrever a terminologia utilizada neste âmbito. Ainda neste Apêndice, caracterizam-se e diferenciam-se os diversos modelos *media* de apoio à decisão.

➤ No *Apêndice C*, uma vez que o sector televisivo pode ser analisado segundo uma grande diversidade de temáticas e perspectivas, que, por sua vez, condicionam as fontes de dados e os métodos a serem adoptados, consta uma síntese dos diversos objectos, disciplinas, orientações e âmbitos de análise que podem ser identificados na investigação televisiva e que viabilizam o enquadramento das análises propostas.

➤ No *Apêndice D* caracterizam-se e comparam-se as perspectivas teóricas que explicam a exposição televisiva, ou seja, a abordagem estrutural e a perspectiva de gratificações e usos do *media*. Apresentam-se os objectivos, os pressupostos, as fontes de dados e as disciplinas associadas a cada vertente de análise, assim como o seu potencial predictivo. Salienta-se a dificuldade de integração de ambas as perspectivas e justifica-se a necessidade de recorrer a análises explicativas da exposição em contextos de uma maior imprevisibilidade dos comportamentos e dos indicadores.

➤ No *Apêndice E* referem-se algumas considerações a respeito da viabilidade das análises de previsão das audiências televisivas e descrevem-se algumas das suas características.

Uma vez que os indicadores de audimetria resultam dos comportamentos de exposição televisiva de um painel, a vertente explicativa e a de previsão encontram-se significativamente relacionadas, apesar de, na bibliografia da especialidade, os indicadores serem, inúmeras vezes, analisados como meras variáveis que apresentam determinada evolução e ignorando os comportamentos que os originam.

Neste Apêndice, caracteriza-se a forte diversidade de contextos de previsão ligados ao planeamento de *media*. Referem-se, também, quais são as características que os sistemas de previsão devem apresentar e a necessidade de complementar diferentes abordagens, assim como a relevância de serem adoptados os métodos quantitativos.

Aborda-se, ainda, o carácter imprevisível dos comportamentos individuais e as dificuldades que as análises de indução apresentam quando o objecto de estudo corresponde a comportamentos humanos. Caracteriza-se o método indutivo e as abordagens metodológicas que lhe estão associadas. As análises de previsão assumem, nas Ciências Sociais, características que as diferenciam das análises que podem ser realizadas nas Ciências Exactas e, neste capítulo, pretende-se contribuir para esclarecer algumas das oposições que existem nestas áreas.

➤ Uma vez que a construção dos modelos de exposição pode, também, implicar e depender da caracterização dos comportamentos de consumo televisivo sob várias perspectivas, no *Capítulo III* consta um resumo dos principais trabalhos desenvolvidos nesta área, sistematizando os vários campos de aplicação das bases de dados de audimetria, incluindo as metodologias utilizadas e os principais resultados alcançados. Este capítulo inicia-se com uma breve introdução à actividade de planeamento de *media* e aos modelos de exposição, descrevendo-se a sua evolução, as metodologias utilizadas e as características que apresentam e que os diferenciam dos modelos que são viáveis nas Ciências Exactas.

No estado da arte, privilegiam-se os modelos explicativos das audiências programáticas e comerciais que estão associados às teorias estruturais da exposição televisiva e os trabalhos de duplicação das audiências. Referem-se os reduzidos contributos teóricos e empíricos ao nível dos modelos explicativos que integram a perspectiva estrutural e a de gratificações e usos do *Media*. Finaliza-se com o estado da arte na área da previsão das audiências programáticas e comerciais, destacando-se as fontes de dados e metodologias utilizadas. Reportam-se os escassos contributos ao nível da análise dos erros de previsão e das perspectivas de previsão que transcendem a previsão dos *ratings*.

➤ A descrição das áreas metodológicas, fontes, bases de dados e amostra consta do *Apêndice A* e *Capítulo IV*. As metodologias, técnicas e algoritmos a utilizar nas análises têm origem em três campos distintos, mas que, nos últimos anos, apresentam uma evolução convergente, nomeadamente a Estatística Multivariada, a Análise Simbólica e a aprendizagem automática ou *Machine Learning*.

Estas metodologias abrangem não só a área da previsão, que se traduz em análises de Regressão linear e não Linear, mas, também, a segmentação de dados, ou seja, a classificação e a análise de agrupamentos e de redução de dimensionalidade dos dados recorrendo a Métodos Factoriais.

➤ No *Capítulo IV* consta uma breve justificação da necessidade de recorrer ao *Data Mining* e à Análise Simbólica e a estrutura metodológica da tese, onde se descrevem os resultados esperados com a aplicação das diversas metodologias. Neste capítulo, consta o enquadramento metodológico e a sequência de abordagens analíticas adoptadas para a realização das análises propostas, referindo-se, para os diversos contextos, o seu carácter inovador face ao que se verifica na literatura da especialidade.

Essencialmente, define-se, nesta parte do trabalho, a linha metodológica mais apropriada a cada objectivo proposto e os potenciais de aplicação de metodologias complementares, algumas das quais não foi possível implementar na prática pela sua inadequabilidade aos dados que foram disponibilizados e pelas restrições temporais, e não só, existentes. Esta tese incide na análise dos modelos de exposição, explorando as potencialidades de áreas estatísticas mais recentes e adequadas à análise das bases de audimetria. Descreve-se, também, no *Capítulo IV*, sinteticamente, a fonte de dados de audimetria e a sequência de bases de dados que foi necessária para construir a amostra. No *apêndice G* comparam-se algumas metodologias.

Do ponto de vista das bases de dados, no *Apêndice A* descreve-se, em maior detalhe, o processo de audimetria, segundo o qual se calculam os índices de audiências, bem como as bases de dados disponibilizadas de índices de audiências programáticos e comerciais. Salienta-se a validade e fiabilidade da audimetria e a sua adequabilidade aos objectivos propostos. Descreve-se, também, a constituição do painel, o processo de amostragem e apresentam-se algumas considerações a respeito da terminologia adoptada. Caracterizam-se os diversos processos de medição de audiências, apresentando-se as respectivas vantagens e limitações.

A apresentação de alguns resultados e a sua discussão constam do *Capítulo V*. Neste capítulo, apenas são referidas as principais conclusões referentes a cada uma das questões que foram colocadas neste trabalho. As análises estatísticas realizadas e uma descrição mais exaustiva dos resultados constam do *Apêndice F*, que se encontra subdividido em sete capítulos e que se descrevem nos seguintes parágrafos.

➤ Nos dois capítulos iniciais, caracterizam-se os três indicadores de audimetria mais relevantes, nomeadamente a evolução e variabilidade observada nas taxas médias de audiências para o total da TV (*Apêndice F.1*), uma vez que este indicador representa a exposição global dos telespectadores ao meio televisivo, e também ao nível dos *ratings*. Por último, de forma muito menos aprofundada, apresentam-se alguns resultados das análises descritivas realizadas aos valores de *shares* de audiências.

Nas análises univariadas, pretendeu-se, também, avaliar a viabilidade de previsão de alguns destes indicadores sem recorrer a variáveis explicativas, bem como o grau da influência sociodemográfica na sua evolução, uma vez que as previsões podem ser realizadas para o universo e, em alternativa, por segmentos, sendo necessário agregar, posteriormente, os resultados (*Apêndice F.2*).

Para desenvolver os modelos de previsão de audiências, poderá ser necessário identificar os factores responsáveis por uma maior variabilidade de valores e identificar os contextos onde se observa uma menor estabilidade.

Em função da evolução e da dimensão dos erros de previsão, será definida a estrutura metodológica mais apropriada. Do ponto de vista da previsão apenas nos contextos de maior instabilidade, justifica-se, porventura, considerar níveis mais desagregados de análise, recorrendo, nomeadamente, aos dados dos ficheiros de base de audimetria, ou, inclusive, a outras fontes de dados que possam contribuir para uma caracterização mais exhaustiva dos padrões de consumo televisivo.

Os *ratings* correspondem ao principal indicador que deverá ser previsto. Contudo, na bibliografia da especialidade, alguns autores (*Napoli, 2001, Patelis, Metaxiotis, Nikolopoulos & Assimakopoulos, 2003*) consideram que será preferível prever os valores das taxas médias de audiências para o total da TV e os valores dos *shares*, resultando os *ratings* do seu produto. Consequentemente, constam, neste Apêndice, análises descritivas dos três principais indicadores nos dois capítulos iniciais. A questão do nível de agregação dos dados a considerar em análises de previsão foi abordada através das análises de segmentação em função da evolução observada nos valores dos *ratings* e das taxas médias de audiências para o total da TV.

➤ Posteriormente, apresentam-se as análises realizadas aos indicadores acumulados. Inicia-se com a análise dos índices de repetição de selecções que permitem avaliar a lealdade programática (*Apêndice F.3*). Segue-se a análise dos índices de duplicação entre programas adjacentes, que viabilizam a análise de um comportamento estrutural designado de efeitos de herança, ou entre programas não adjacentes e entre faixas horárias, assim como aos fluxos de audiências entre canais (*Apêndice F.4*). Por último, consta a análise da lealdade ao canal e tipológica (*Apêndice F.5*).

As análises de duplicação das audiências têm como função avaliar se existe regularidade na exposição do painel ao *media* e aos suportes. Pretende caracterizar-se alguns comportamentos de consumo televisivo e avaliar-se se a previsão será viável, recorrendo a determinadas metodologias de *Data Mining* que, em função dos resultados obtidos, são propostas no *Capítulo VI*.

No *Apêndice F.5* constata-se se, recorrendo às variáveis sociodemográficas que constam das bases de audimetria, é possível identificar, para cada um dos segmentos, se existe uma maior exposição a suportes com determinadas características. Em determinadas etapas intermédias de construção dos planos de *media* existe a necessidade de restringir e seleccionar os suportes mais adequados ao público objectivo. Avalia-se em que medida determinadas abordagens metodológicas permitem a avaliação do perfil das audiências dos suportes e a sua selecção, assim como a segmentação do painel.

Realizaram-se algumas análises aos diversos indicadores sociodemográficos que permitem a avaliação de alguns comportamentos, nomeadamente da fidelidade programática (*Apêndice F.3*) e da lealdade tipológica e ao canal através da identificação dos conteúdos que mais se associam às diversas variáveis de caracterização do painel. A previsão de alguns indicadores de audimetria não depende apenas da quantificação do número de telespectadores que se encontram expostos aos suportes, mas, também, da duração dessa exposição. Nesse sentido, a análise da fidelidade programática permite avaliar qual será o grau de permanência dos telespectadores no decorrer dos programas e quais são os factores que o condicionam. A questão da segmentação das audiências será desenvolvida em maior grau no *Capítulo VI*.

➤ Nas análises realizadas aos *ratings* programáticos recorre-se a metodologias lineares e não lineares multivariadas (*Apêndice F.6*), que visam o desenvolvimento de modelos explicativos e, por vezes, de previsão dos seus valores, considerando os factores explicativos estruturais que constam das bases de audimetria. Esta tipologia de dados deve ser privilegiada, uma vez que as agências de meios, na sua maioria, apenas têm acesso aos dados de *ratings*.

Verifica-se se estas análises conduzem a modelos de previsão com maiores percentagens de variância explicada face aos resultados das análises univariadas realizadas aos valores dos *ratings* e que constam do *Apêndice F.2*. Neste capítulo, avalia-se se as análises não lineares apresentam menores erros de previsão, recorrendo a bases de dados nacionais e a uma amostra que considera dados a partir do ano de 2000, e que corresponde a uma lacuna existente na bibliografia nacional e internacional da especialidade.

Uma vez que a Regressão linear corresponde à metodologia utilizada com uma maior frequência na bibliografia internacional, avalia-se, também, em que medida os algoritmos simbólicos permitem que melhores resultados sejam alcançados. As análises de previsão devem ser de fácil utilização, exequíveis em tempo útil, que possam aplicar-se a diferentes públicos alvo e que viabilizem a baixo custo uma previsão a curto e médio prazo. Estes factores foram considerados quando se seleccionaram as abordagens metodológicas, algumas das quais foram implementadas neste capítulo, sendo outras sugeridas no *Capítulo VI*.

➤ No que diz respeito às audiências comerciais (*Apêndice F.7*), pretende-se, também, o desenvolvimento de análises explicativas lineares e não lineares de *ratings* que permitam a avaliação da repercussão dos factores estruturais que constam das bases de audimetria nos valores dos índices de audiências.

Simultaneamente, realizaram-se análises a um indicador que foi calculado e designado por *Spot_Ratio%*, que quantifica a inter-relação existente entre as audiências programáticas e as comerciais, resultando do seu quociente. Avalia-se o grau de previsibilidade de ambos os indicadores, recorrendo a análises lineares e não lineares e diferenciando as análises em função dos canais e de determinadas faixas horárias. Finaliza-se com a avaliação da influência de algumas variáveis sociodemográficas na exposição aos espaços comerciais.

No *Capítulo V* consta uma síntese dos principais resultados, que foi complementada com o *Apêndice F*. A discussão de resultados, a nível substantivo e metodológico, consta do *Capítulo V*, no qual se avaliam, sucintamente, as questões apresentadas no *Capítulo I.3* e no *Capítulo II*.

Neste capítulo constam algumas considerações a respeito das melhores abordagens metodológicas implementadas no âmbito da previsão dos *ratings* para diversos contextos e que, posteriormente, serão desenvolvidas através de uma proposta metodológica que possa contribuir para que melhores resultados sejam alcançados na área do planeamento de *media*. Apresenta-se e caracteriza-se um modelo, essencialmente estrutural e explicativo da exposição televisiva, o qual permite compreender que o consumo televisivo transcende as “preferências” do público e que a multiplicidade de factores e as suas características impedem que bons resultados sejam alcançados quando se pretende promover a previsão de comportamentos de consumo televisivo.

Atendendo aos resultados das análises realizadas, no *Capítulo VI* consta a descrição da sequência de análises sugerida para viabilizar a construção dos modelos de exposição em função dos contextos, isto é, quer ao nível da previsão de alguns índices de audiências, com as menores margens de erro possível, quer quando perspectivas de análise alternativas se mostram necessárias. A sequência de análises a adoptar será função do grau de variabilidade que se observa nos principais indicadores ou nos restantes dados de audimetria e da dimensão da repercussão da programação concorrente e adjacente.

As análises de previsão podem resultar de análises univariadas ou multivariadas das variáveis e indicadores que constam das bases de dados de audimetria do ponto de vista agregado ou, eventualmente, implicar a necessidade de serem complementadas com outras fontes de dados que caracterizem, com uma maior profundidade, os factores que condicionam a exposição de cada agregado familiar e de cada indivíduo.

➤ As conclusões finais, o sumário de contribuições, as limitações das análises realizadas e as propostas para análises futuras constam do *Capítulo VII*. As referências bibliográficas constam do *Capítulo VIII*.

A *Figura I.5.1* ilustra a estrutura descrita anteriormente, e na qual os *Apêndices B* até *E* complementam o estado da arte e têm como função introduzir a temática do planeamento de *media* e dos modelos explicativos e predictivos nas Ciências Sociais e na área da audimetria. Proporciona-se, também, uma perspectiva global da investigação realizada neste sector, facilitando, deste modo, o enquadramento das análises propostas.

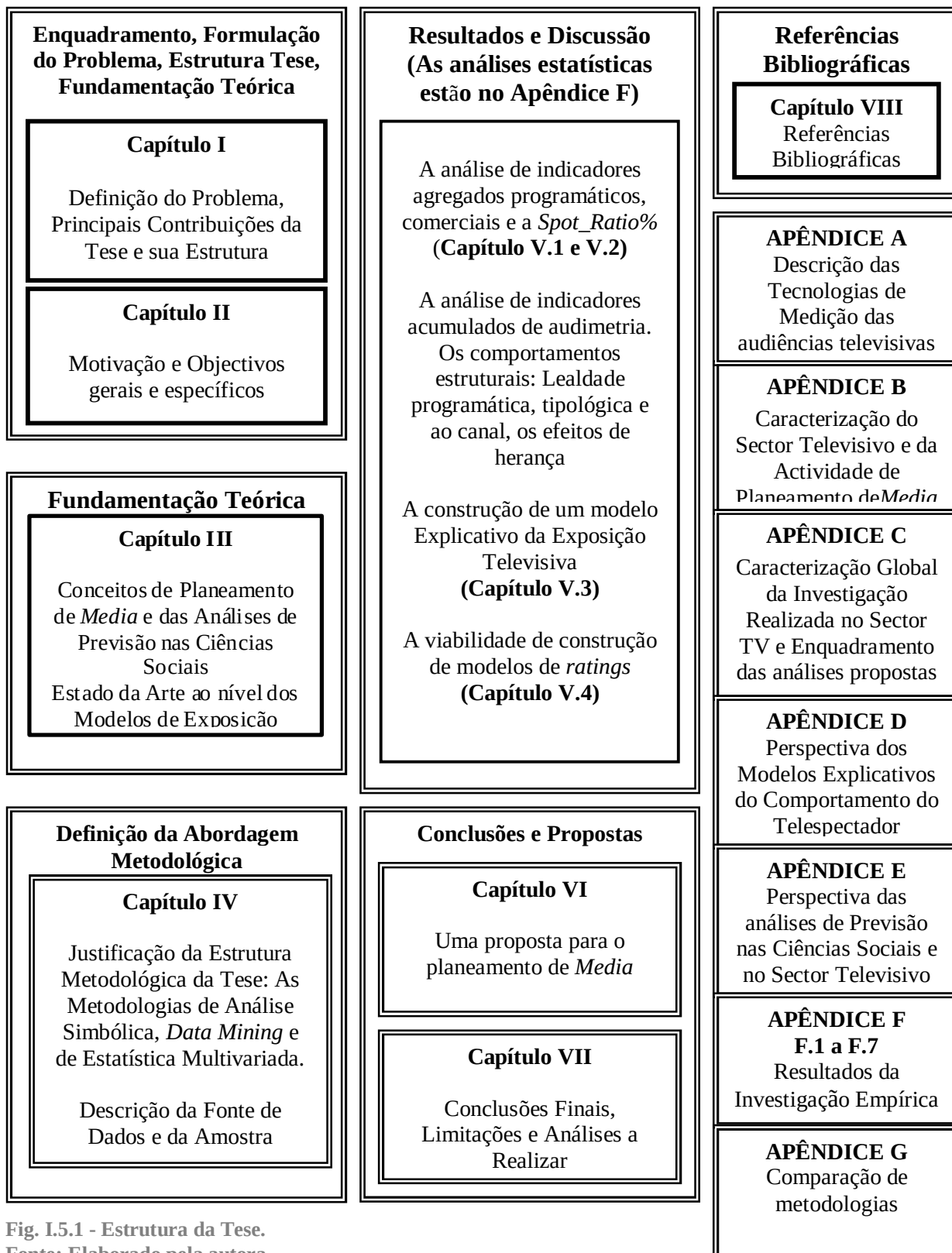


Fig. I.5.1 - Estrutura da Tese.
Fonte: Elaborado pela autora.

CAPÍTULO II _____

OBJECTIVOS E QUESTÕES A DESENVOLVER

Este capítulo faz referência ao objectivo geral e aos objectivos específicos da presente tese em função das fontes de dados e dos restantes recursos disponíveis para a sua realização. O objectivo geral traduz-se na identificação das melhores perspectivas de análise quando se pretende a construção dos modelos de exposição. Os objectivos específicos representam o conjunto de análises que possibilitam compreender a evolução dos dados e avaliar a viabilidade de construção dos modelos de “*ratings*”. Elucidam-se as quatro questões a desenvolver, permitindo, assim, uma maior percepção acerca da sequência de resultados apresentados. A conclusão aborda a descrição das questões operacionais em análise.

II.1. MOTIVAÇÃO E OBJECTIVO GERAL

Na actualidade, os estudos e as investigações multidisciplinares proliferam na área televisiva, oferecendo a sua particular visão deste sector. A necessidade de seleccionar determinado aspecto, implica a escolha de uma parte específica do estudo e assumir um determinado ponto de vista, condicionando as abordagens analíticas e as fontes de dados necessárias e adequadas aos objectivos propostos.

As análises a realizar pretendem contribuir para a resolução de questões que são relevantes no âmbito do planeamento de *media*, condicionando o desempenho organizacional. Visam o desenvolvimento dos modelos de exposição que permitem compreender e prever as audiências televisivas, sendo inevitável recorrer à audimetria.

Incidem em questões que preocupam os profissionais deste sector e pretendem contribuir para a construção de determinados modelos de *media*, os quais apresentam uma crescente complexidade e colocam inúmeras dificuldades a quem pretende investir em publicidade televisiva.

Simultaneamente, as Ciências Estatísticas contribuem para o estudo do sector televisivo através dos seus próprios instrumentos que tornam viável retirar o máximo de informação dos dados recolhidos. Os resultados obtidos dependem muito das abordagens metodológicas adoptadas no estudo. Consequentemente, privilegiou-se a avaliação do contributo de áreas metodológicas mais recentes, e adequadas, à análise das bases de dados de audimetria que apresentam uma elevada dimensionalidade. Na bibliografia da especialidade predominam abordagens e perspectivas, frequentemente pouco adequadas à sua análise e, sem terem em atenção as limitações e restrições que as análises de previsão apresentam nas ciências sociais.

O desempenho de inúmeras empresas ligadas ao sector televisivo encontra-se dependente da previsão e do desenvolvimento de modelos explicativos da evolução de determinados indicadores de audimetria. As estações televisivas definem a estrutura da grelha de programas e as tarifas publicitárias em função dos valores dos índices de audiência dos programas, faixas horárias e espaços comerciais. Os anunciantes e as agências de meios necessitam por isso de conhecer qual será o número de telespectadores expostos aos suportes que pretendem adquirir, os quais devem ser os mais adequados à divulgação da mensagem publicitária. A sua selecção é função da frequência de exposição e da sua cobertura do mercado alvo, sendo, também, fundamental conhecer a composição sociodemográfica da sua audiência.

O objectivo geral está representado na seguinte questão “Qual o contributo das metodologias de Data Mining e de Análise Simbólica para o desenvolvimento de modelos de “ratings”, enquanto tipologia de dados associada ao processo de audimetria que contribuam para a construção dos modelos de exposição que são essenciais para a tomada de decisão a nível operacional e estratégico no sector televisivo?”

Os modelos explicativos, e de previsão de valores de *ratings* programáticos e comerciais, correspondem a uma área crucial para a gestão das empresas ligadas a este sector de actividade, sendo esse o objectivo deste trabalho. São ainda avaliadas as influências estruturais na evolução dos indicadores, e a capacidade explicativa e predictiva de diferentes abordagens metodológicas, enquadradas na Análise Simbólica e no *Data Mining*. Os modelos de “*ratings*” devem caracterizar o comportamento agregado, quer a nível programático, quer a nível comercial, atendendo à sua interdependência que apenas, pontualmente, tem sido mencionada na bibliografia da especialidade.

Previamente será necessário conhecer a evolução dos dados a diversos níveis, ou seja, desenvolver modelos explicativos das variáveis que condicionam a construção dos planos de *media* e que não constam da bibliografia nacional da especialidade, nomeadamente dos indicadores agregados e acumulados. As análises permitem, conjuntamente, retirar ilações quanto à forma do painel interagir com o *media* e os diversos suportes. Devemos identificar os contextos que podem requerer a aplicação de diferentes abordagens metodológicas do ponto de vista predictivo. Conseguir identificar as alternativas de análise que viabilizem a construção dos planos de *media*, será função do perfil de exposição ao *media* do painel de audimetria, e da evolução dos dados que dela resulta. Em função dos resultados das análises exploratórias e descritivas, a viabilidade dos modelos de *ratings* deverá ser avaliada e as perspectivas alternativas de análise deverão ser identificadas, de modo a permitir que contribuam para a resolução de questões que implicam a necessidade de prever e compreender a exposição televisiva e a evolução dos fluxos de audiências.

Em suma, pretende-se identificar as melhores perspectivas de construção dos modelos de exposição, considerando metodologias inexistentes ou que, apenas pontualmente, tenham sido avaliadas por outros autores, considerando um conjunto de variáveis explicativas e uma amostra muito mais abrangente. As abordagens metodológicas condicionam fortemente os resultados que se podem obter, e, segundo a bibliografia da especialidade, imperam as mais tradicionais, em especial as análises lineares.

II.2. OBJECTIVOS ESPECÍFICOS

As análises de previsão nas Ciências Sociais apresentam grandes limitações e, consequentemente podem tornar inviável a obtenção de bons resultados aquando da análise dos índices de audiência, ou dos comportamentos de consumo televisivo de um painel de audimetria. Um maior entendimento sobre as variáveis que condicionam os “*ratings*” pode contribuir para a maximização dos seus valores e para uma correcta estruturação das grelhas de programação, assim como para a identificação das dimensões responsáveis pelas dificuldades que o sector enfrenta, quando tem que realizar as análises de previsão.

Os modelos de *ratings* foram privilegiados porque correspondem à tipologia de dados mais acessível aos profissionais do meio e, que lhes permitirá alcançar uma solução da forma mais rápida, simples e menos onerosa. O planeamento de *media* depende do comportamento global do painel. Segundo a bibliografia da especialidade, existe uma maior repercussão das variáveis estruturais na exposição televisiva a nível individual, sendo necessário avaliar qual será a sua repercussão na exposição agregada e na evolução dos principais indicadores de audimetria. O comportamento agregado apresenta características específicas e, por vezes, não corresponde ao que se verifica a um nível mais desagregado.

Pretende-se analisar os principais indicadores agregados que permitem otimizar o investimento em publicidade televisiva, nomeadamente dos valores de *ratings*, taxas médias de audiências para o total da TV e dos *shares* de audiência. Privilegia-se a análise dos indicadores associados à globalidade do painel, uma vez que os indicadores universais correspondem aos mais relevantes para a construção das campanhas publicitárias. Regra geral, face ao carácter mais oneroso da publicidade televisiva, a maior parte das empresas que recorrem a este *media* pretendem que os seus anúncios estejam expostos ao maior número de telespectadores, dos mais diversos segmentos. Apenas para produtos dirigidos a determinados nichos de mercado, poderá ser relevante considerar um conjunto mais restrito de telespectadores.

Os indicadores acumulados, também, são essenciais neste processo, uma vez que condicionam a selecção dos suportes e permitem interpretar a “qualidade” dos *ratings*. A análise dos comportamentos estruturais e dos indicadores de duplicação das audiências, possibilita-nos compreender como o painel interage maioritariamente com o *media*, o que, também, condicionará as abordagens metodológicas mais adequadas para a futura quantificação das audiências e para a identificação dos suportes nos quais investir. Devemos avaliar qual é o grau de regularidade da exposição e a sua repercussão na evolução dos *ratings* e na viabilidade da sua previsão.

As análises a realizar aos índices de duplicação, permitem caracterizar os hábitos dos telespectadores ao nível dos programas ou das faixas horárias, onde vão estar inseridas as inserções publicitárias. Quando se constroem planos de *media*, os suportes, também, são seleccionados em função do grau de “repetição” da sua exposição aos telespectadores que integram o público-alvo. Um anúncio, deverá estar exposto ao mesmo telespectador, um determinado número de vezes, para o conseguir influenciar e deste modo favorecer a aquisição dos produtos divulgados.

A taxa média de audiência de um programa, também, depende dos comportamentos de lealdade programática, efeitos de herança entre programas, fluxos de audiências entre os canais e dos comportamentos de lealdade tipológica e ao canal. Através da avaliação destes factores, contribui-se, também, indirectamente, para a caracterização de alguns comportamentos de consumo televisivo. Permite-nos avaliar qual é o grau de passividade e regularidade inerente à exposição televisiva. Em função da existência de uma maior actividade ou passividade subjacente às decisões de acesso ao *media* e de selecção dos programas, poderá ser necessário adoptar diferentes abordagens, que tornem viável o cálculo dos futuros valores dos principais indicadores. Ainda que os indicadores agregados e acumulados para o universo televisivo sejam os mais relevantes, por vezes, outros perfis de audiência, podem ser seleccionados pelos anunciantes.

A maioria dos trabalhos publicados no âmbito dos índices de duplicação e das análises explicativas e previsão de índices de audiências, ignora totalmente a repercussão das variáveis sócio-culturais na sua evolução. Do ponto de vista do planeamento de *media*, os indicadores universais são, também, os mais relevantes. Porém, do ponto de vista da explicação e previsão da evolução dos dados, considerando a totalidade do painel, também, é relevante avaliar quais são os segmentos que podem incutir uma maior variabilidade, ou estabilidade, à sua evolução global. Os segmentos mais instáveis devem ser analisados posteriormente, de forma mais aprofundada, para se alcançar uma melhor percepção de como estes interagem com este *media*.

Uma das questões que é, também, proeminente para a planificação das campanhas publicitárias, consiste na redução da dimensionalidade do número de suportes. Considerando os indicadores sociodemográficos que constam das bases de audimetria, podemos caracterizar o perfil das audiências programáticas, tipológicas e dos diversos canais e, na sua sequência, associar segmentos aos suportes publicitários. A dimensão das bases de dados de audimetria, a sua própria estrutura e o grau de dispersão dos dados, não viabiliza que as análises sejam conduzidas para a totalidade dos segmentos existentes. Os trabalhos publicados sobre a análise dos índices de duplicação incidem nas faixas horárias associadas aos períodos de maiores audiências; consideram os dados associados ao universo televisivo e recorrem a metodologias mais tradicionais. Ao nível das análises de previsão e explicativas da evolução dos indicadores acumulados, a maioria dos trabalhos debruça-se na apresentação de diferentes propostas de abordagens metodológicas, considerando apenas os dados para o universo televisivo e algumas escassas variáveis estruturais.

Na presente tese, considera-se a totalidade da emissão diária nas análises a realizar; recorre-se às áreas metodológicas mais recentes e em alguns contextos, também, se avalia a repercussão das variáveis sociodemográficas. A profundidade das análises, a este nível, encontra-se condicionada pelos recursos disponíveis e por parte do investimento temporal necessário para a estruturação das mesmas.

Uma vez que não está em causa a análise de percepções, motivações e atitudes de consumo televisivo, os dados de audimetria devem ser considerados a um nível mais desagregado, apenas quando for superior a variabilidade dos indicadores calculados para o universo televisivo. No entanto, avaliou-se sempre que possível, a repercussão das variáveis sociodemográficas na evolução dos indicadores, nomeadamente, ao nível das taxas médias para o total da TV, *ratings* programáticos, índices de fidelidade, afinidade, atractividade e de adesão. Não foi possível identificar, na literatura da especialidade, análises quantitativas que tenham como objectivo a caracterização destes indicadores.

Do ponto de vista explicativo da exposição televisiva, e dos modelos de *ratings*, persistem inúmeras incógnitas. Desconhece-se a capacidade explicativa e de previsão de inúmeras variáveis estruturais e individuais, assim como as interacções existentes. Também se desconhece quais são os segmentos existentes e a sua caracterização, assim como o poder discriminante das variáveis sociodemográficas, entre outras. As abordagens metodológicas que permitem retirar uma maior informação e conhecimento, a partir das bases de audimetria, não estão identificadas. Consequentemente, as análises realizadas, também pretendem contribuir parcialmente para a resolução de algumas destas questões.

Como em qualquer outro estudo quantitativo, promover a previsão das variáveis mencionadas, ou de qualquer outro indicador, requer em primeiro lugar a caracterização da sua evolução, nomeadamente, avaliando qual é a sua variabilidade e a identificação dos factores com uma maior repercussão a este nível, constatando-se quais os factores controláveis, que podem ser calculados com algum rigor no momento de realizar uma análise de previsão, ou não. Devemos avaliar qual deverá ser o nível mais adequado de agregação dos dados. Em seguida, definem-se as abordagens metodológicas mais adequadas em função dos contextos, por exemplo, seleccionar entre os métodos lineares ou não lineares, univariados ou multivariados, e do âmbito da estatística mais tradicional ou enquadrados em áreas mais recentes, como por exemplo, na Análise Simbólica e no *Data Mining*.

A presente tese promove a aplicação de metodologias de Análise Simbólica e de *Data Mining* em contextos que não constam da bibliografia da especialidade, a qual foi identificada e com uma forte diversificação ao nível dos métodos analíticos adoptados. Recorre-se a uma fonte de dados que a nível Europeu é menos frequente e, realizaram-se algumas análises que, por vezes, surgem na bibliografia internacional, mas não com base em dados nacionais. Porém, adoptando uma nova perspectiva face ao que se verifica a nível internacional e tendo em consideração novas variáveis, e abordagens metodológicas, mais adequadas à sua análise.

A necessidade de diferenciar as análises explicativas das análises de previsão, decorre de frequentemente se classificar como análises de previsão os trabalhos que procuram identificar quais as variáveis que condicionam determinado resultado, mas sem conseguir prever observações concretas no espaço e no tempo. Por vezes, as variáveis que são identificadas, também, não são conhecidas *a priori*, nem são previsíveis. Uma variável pode condicionar a probabilidade de se verificar determinado resultado, mas, desconhecendo-se quando este irá ocorrer, tendo como exemplo algumas variáveis que influenciam a aquisição de determinado produto ou serviço.

Os modelos que promovem a previsão de audiências televisivas distinguem-se em modelos explicativos ou univariados e em lineares, não lineares ou híbridos. A segunda classificação é efectuada em função das áreas metodológicas adoptadas. Em relação à categorização inicial, os modelos univariados não consideram variáveis exógenas, apenas desfasamentos temporais entre as observações da variável a prever. A grande maioria corresponde a modelos sazonais univariados, desenvolvidos por aplicação de metodologias de análise de séries temporais e, que calculam as previsões com base na análise da sazonalidade associada ao consumo histórico, negligenciando outras variáveis que podem ser fundamentais para a sua previsão. As previsões que recorrem às séries temporais têm uma relativa importância na investigação televisiva e na optimização do planeamento comercial.

Os métodos clássicos de previsão de audiências com base na análise de séries temporais, por exemplo, as análises ARIMA e SARIMA, Médias Móveis ou análises de Fourier, têm como vantagens a sua simplicidade e a disponibilidade de métodos para controle do sobreajustamento. Todavia, uma vez que podem negligenciar variáveis importantes para a previsão, não é viável uma interpretação com significado dos resultados alcançados e, consequentemente estes apresentam uma menor relevância prática.

O que se pretende neste estudo corresponde, essencialmente, ao desenvolvimento de análises explicativas lineares e não lineares, em detrimento de meras previsões univariadas de audiências programáticas e comerciais. Considerando variáveis exógenas, os resultados obtidos podem se manter válidos por um horizonte temporal maior. Privilegiam-se os modelos que incluem variáveis explicativas estruturais, em detrimento da análise das séries temporais multivariadas.

Os modelos lineares de *ratings* que constam da bibliografia da especialidade, não têm frequentemente permitido bons resultados. Uma vez que os contextos socioeconómicos e culturais, para além da estruturação do mercado influenciam os resultados, também, será necessário avaliar a viabilidade de considerar os modelos de Regressão linear, cuja capacidade explicativa serve de base de comparação face aos restantes métodos implementados. As análises lineares, resultam da aplicação de metodologias de Regressão simbólica e não simbólica. As análises não lineares foram realizadas recorrendo a metodologias da área do *Data Mining* Proposicional, Temporal e Simbólico.

Ao nível do *Data Mining* privilegiam-se as metodologias proposicionais e algumas metodologias temporais ou sequenciais, mas em função dos resultados alcançados, referindo e explorando as potencialidades de aplicação de metodologias das restantes áreas do *Data Mining*, nomeadamente de DM estruturado, ou mais intensivas do ponto de vista computacional no âmbito Temporal e Sequencial.

Pontualmente, também, se avalia a viabilidade de previsão de alguns dos indicadores mais relevantes para o planeamento de *media*, ou descrevem-se as abordagens metodológicas que possam contribuir para o efeito. Por vezes, os reduzidos coeficientes de variância explicada, que são apresentados na literatura da especialidade, resultam de uma codificação inapropriada das variáveis em estudo, de uma incorrecta avaliação das interacções existentes e da inadequação das metodologias seleccionadas, para que possam ser atingidos os objectivos propostos. Acabam por ser tirar conclusões erróneas e de se verificar uma subutilização das bases de dados disponíveis.

Se as análises univariadas e multivariadas dos principais indicadores de audimetria não apresentarem razoáveis margens de variância explicada, a previsão não será viável sem a interpretação da estrutura dos comportamentos de consumo televisivo, podendo-se associar probabilidades a determinados cenários de exposição uma vez que o comportamento humano é estocástico. Porém, uma caracterização com uma maior profundidade da exposição, não será viável sem complementar as análises, recorrendo a outras fontes de dados que permitam considerar um conjunto muito mais abrangente de variáveis de caracterização da amostra, do processo de selecção programática e de exposição ao *media*. Devem ser avaliadas alternativas de análise nos contextos de maior variabilidade dos valores de *ratings*.

Uma vez que as análises que contam da literatura da especialidade incidem, maioritariamente, na estatística mais tradicional, salientam-se os diversos potenciais de aplicação de áreas de análise estatística mais recentes, não só para a previsão dos indicadores de audimetria, mas também, para retirar informação relevante para o sector a partir das bases de audimetria e, relativa, às diferentes formas de interacção da audiência com o *media*. Os factos identificados podem ser explicados e aprofundados, se necessário, por parte das disciplinas aptas a realizá-lo, na eventualidade de ser necessário recorrer à investigação qualitativa. Demonstra-se nos seguintes capítulos, a necessidade de diversificação dos métodos de análise e o reconhecimento de que pela via da integração e multidisciplinariedade melhores resultados podem ser alcançados.

Encontra-se bem documentado na literatura da especialidade as limitações de recorrer exclusivamente a abordagens mais qualitativas, inseridas nas abordagens das gratificações e usos do *Media*, ou por aplicação de outros métodos de medição das audiências, não só para a previsão de indicadores que são relevantes para a sobrevivência das organizações, como também para uma melhor percepção e caracterização do consumo e da evolução dos índices de audiências televisivas. Quando se pretendem realizar análises parciais e não integrativas das diversas perspectivas de caracterização da exposição televisiva, as teorias estruturais do consumo televisivo, têm conduzido aos melhores resultados no decorrer das últimas décadas, tendo em atenção a futura quantificação da exposição televisiva.

O consumo televisivo encontra-se condicionado por inúmeras variáveis, as quais interagem entre si de forma variável no decorrer do tempo, inúmeras das quais, por vezes, dificilmente quantificáveis ou previstas, com reduzidas margens de erro. Os modelos explicativos podem ser considerados como análises preliminares e subjacentes ao desenvolvimento de modelos de previsão.

O objectivo geral traduz-se, por exemplo, na *“avaliação das seguintes quatro questões que abordam o desenvolvimento de modelos explicativos e de previsão de “ratings”*:

Questão 1: “Qual o poder explicativo dos factores estruturais e das variáveis individuais que constam das bases de dados de audimetria ao nível dos “ratings” programáticos, comerciais e dos “factores bloco”, que representam as quedas observadas nas audiências no decorrer dos espaços comerciais?”

Pretende-se avaliar que “a evolução dos “ratings” depende maioritariamente dos “ratings, avaliando sempre que possível as interacções existentes entre as variáveis explicativas. Quando as análises explicativas identificam factores conhecidos *a priori*, ou que sejam previsíveis, estes podem ser utilizados posteriormente em modelos de previsão do indicador em análise.

Questão 2: “Qual o modelo estrutural explicativo que permite caracterizar maioritariamente a exposição televisiva do painel de audimetria e de que forma condiciona a viabilidade da quantificação da exposição televisiva?”

A análise dos “comportamentos estruturais”, tem como intuito caracterizar a exposição televisiva segundo determinadas dimensões, as quais condicionam a venda do espaço comercial. Apresenta como objectivo final, avaliar se a audiência é variável quando se relacionam os indivíduos e os suportes, uma vez que a dinâmica dos comportamentos condiciona a viabilidade da previsão da exposição e a componente metodológica a adoptar. Avalia-se a predisposição para se desenvolverem determinados comportamentos, respectiva interacção e factores que os condicionam, e, por último a sua repercussão na evolução dos “ratings”.

Questão 3: “Qual a viabilidade de construção de planos de *media*, antes de serem aplicados os processo de optimização, com base na previsão de valores de “ratings”, recorrendo a metodologias multivariadas, simbólicas e de *Data Mining*, assim como para a aquisição de suportes a nível individual?”

Deve ser privilegiada a perspectiva de análise dos *ratings*, uma vez que, corresponde à única tipologia de dados à qual a maioria das agências de meios, e anunciantes, tem acesso. Os planos também podem ser construídos, considerando a cobertura e a frequência média de exposição. Consequentemente, será necessário avaliar qual é a viabilidade de construção de planos de *media* por previsão dos fluxos de audiências entre os suportes.

Questão 4: “Em função das análises explicativas realizadas e dos resultados das análises de previsão realizadas aos valores dos “ratings” quais são as abordagens metodológicas e os recursos alternativos que permitem obter melhores resultados quando se pretende desenvolver um plano de *media*, uma vez que a maior parte das receitas publicitárias decorrem das vendas realizadas no “up-front market”?”

II.3. AS QUESTÕES OPERACIONAIS EM ANÁLISE

A ausência ou escassez de publicações nas principais revistas da especialidade referentes à análise das bases de dados de audimetria nacionais não facilita a definição de “hipóteses” e questões a desenvolver, que permitam confirmar ou contestar os resultados anteriormente divulgados por outros autores, ao contrário do que se verifica em países anglo-saxónicos. As questões a desenvolver em cada um dos capítulos foram definidas com base nas análises necessárias para conseguir compreender a evolução dos dados, e na sua sequência, permitir a identificação das perspectivas de análise mais adequadas à construção dos modelos de exposição e às restantes necessidades de informação que surgem no planeamento de *media*.

A *Figura II.3.1* representa alguns dos factores e contextos que, talvez, possam influenciar em maior grau a exposição programática e comercial dos telespectadores e a evolução dos indicadores que dela resultam.

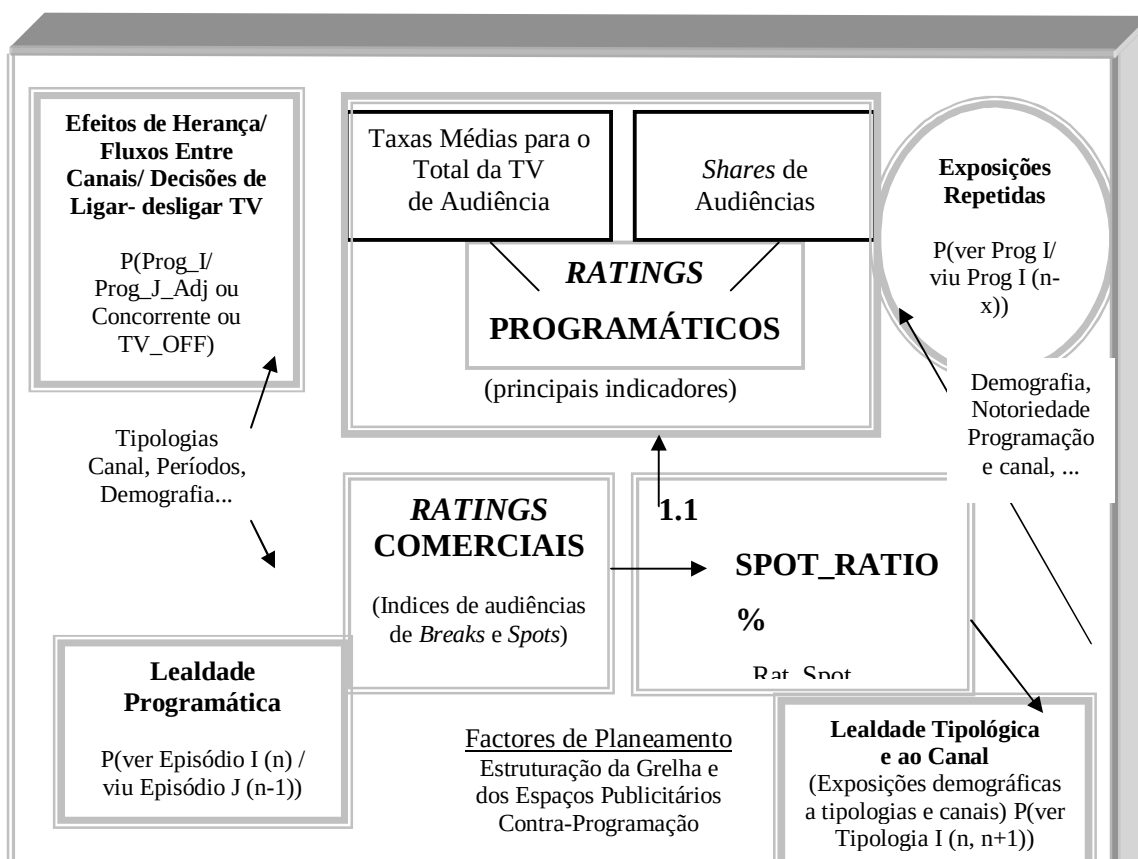


Fig. II.3.1 - Factores que condicionam os índices de audiências e a actividade de planeamento de *Media*. Fonte: Elaborado pela autora.

A análise dos itens representados na *Figura II.3.1*, deverá permitir caracterizar a evolução dos principais indicadores, assim como identificar os factores que os condicionam, avaliando sempre que possível as respectivas interdependências.

Tradicionalmente, a previsão dos indicadores de audimetria era realizada partindo do pressuposto de que o futuro será igual ao passado recente, ou que o erro que se comete segundo esta perspectiva será reduzido, baseando-se os cálculos apenas na média dos resultados históricos para as faixas horárias equivalentes. Esta perspectiva de análise também foi avaliada, uma vez que a nível nacional ainda persiste no dia-a-dia de determinadas agências de meios. Frequentemente, as análises de previsão recorrem apenas às médias de valores históricos para idênticos períodos horários ou de valores de *ratings* de programas anteriormente transmitidos em circunstâncias consideradas análogas, ou que apresentem grande semelhança com a observação a prever.

Devemos verificar em que medida as faixas horárias, no decorrer dos anos, apresentam um equivalente grau de variabilidade e evolução pertencendo aos mesmos segmentos horários. Se eventualmente não existir a mesma classificação de faixas horárias, podemos concluir que recorrer à sua média não será certamente a melhor opção.

A divisão da emissão diária em segmentos horários com evoluções comparáveis dos indicadores em estudo, também corresponde a um dos factores mais relevantes e que, deverá ser incluído, em análises de previsão uma vez que condicionam não só as abordagens metodológicas mais adequadas, mas por vezes, a própria classificação horária dos programas pode ser considerada como variável explicativa ou dependente nas análises de previsão. Por exemplo, em análises de identificação de cenários comparáveis à observação a prever, a classificação horária é determinante. Quando se consideram programas de períodos horários que se diferenciam significativamente, têm que ser realizados os necessários ajustamentos quando se ponderam os valores dos seus *ratings*, no sentido de se prever o valor do indicador de audimetria para uma nova observação.

Nas faixas horárias que apresentam uma maior variabilidade, as previsões são realizadas por vezes com base em sondagens de intenção de ver, ou em juízos de valor de um painel que transmitem as suas percepções, relativamente à evolução mais provável dos indicadores de audimetria. Logo, as análises descritivas e explicativas apresentam uma grande relevância, permitindo avaliar as circunstâncias em que se verifica uma menor regularidade na exposição televisiva e a identificação dos factores que condicionam em maior grau a variabilidade dos valores de *ratings*.

Em determinadas faixas horárias, a combinação de vários métodos, poderá tornar viáveis os melhores resultados, sendo necessário diferenciar os horizontes temporais em função da evolução dos indicadores em estudo. Quando a programação não varia substancialmente, assim como a evolução das variáveis em estudo, não há necessidade de utilizar modelos mais complexos de previsão, meras extrapolações de valores são suficientes para prever os resultados necessários à construção dos planos de *media*.

Devemos ter em atenção que a previsão final será realizada pelo analista, em função da sua experiência e conhecimento do *media*, para além da sua intuição. As análises de previsão devem ser simples de realizar e rápidas uma vez que, apenas correspondem a um valor utilizado como estimativa inicial e sobre o qual, posteriormente, são realizados os ajustamentos considerados necessários. Para que essas alterações possam ser realizadas, será necessário perceber o funcionamento dos modelos e, assim, conseguir diferenciar o que é erróneo, ou não. Não será vantajoso dispor de um modelo que, ocasionalmente proporciona resultados exactos, mas que a maioria das vezes, poderá produzir um fluxo duvidoso de dados, não susceptíveis de, no final, serem ajustados para melhorar o seu grau de precisão.

Existindo estabilidade nos dados, as previsões podem ser realizadas com base na análise das séries temporais, projectando-se as tendências observadas e calculando-se os seus coeficientes de estacionalidade.

Os factores estacionais devem ser calculados separadamente por mês, faixas horárias e por público-objectivo. Uma vez que, as análises de previsão podem ser realizadas para determinados segmentos, também deverá ser comparada a evolução dos indicadores de audimetria para os diversos segmentos representados nas bases de dados, discriminando os resultados em função dos canais, faixas horárias e dias.

As análises a realizar incidem num período característico, a nível médio, do consumo televisivo, por conseguinte, não faz sentido, falar em tendências e em coeficientes de estacionalidade. Pretende-se sim, avaliar a viabilidade de recorrer a várias metodologias de Regressão não linear, para tentar prever os valores das taxas médias de audiências programáticas e para o total da TV. Também deverá ser avaliado, em que medida a análise de séries temporais recorrendo a metodologias mais recentes permitem realizar previsões com menores margens de erro, considerando, não só, metodologias de previsão, mas também de segmentação e *clustering* dos dados.

Com a desregulamentação do sector televisivo, e um acréscimo ao nível da oferta televisiva, as análises de previsão passaram a incluir algumas variáveis explicativas, nomeadamente, as variáveis temporais, o canal e as tipologias transmitidas. Verifica-se a necessidade de avaliar os modelos multivariados de *ratings*, considerando a influência de inúmeros factores estruturais, quer do ponto de vista linear, quer em modelos não lineares e, recorrendo a áreas metodológicas mais recentes. Os modelos de *ratings* devem ser definidos para os *ratings* comerciais, programáticos e para um indicador designado de *Spot_Ratio%*, que inter-relaciona ambas as vertentes. Na realidade, as audiências comerciais, correspondem ao indicador que é mais relevante para este sector, contudo, a sua previsão, encontra-se dependente do respectivo enquadramento programático, apesar das suas audiências apresentarem a sua própria dinâmica.

A previsão das audiências poderá ser realizada através da avaliação dos comportamentos de duplicação de audiências. Na literatura da especialidade, as diferentes vertentes de análise dos comportamentos estruturais, são avaliadas de forma isolada e sem ter em atenção o seu contributo ao nível da previsão dos indicadores de audimetria. As análises incidem em maior grau ao nível dos efeitos de herança entre programas directamente adjacentes e são escassas as publicações que apresentam como objecto de estudo as restantes tipologias de comportamentos. Devem ser avaliadas as abordagens metodológicas mais adequadas para a sua análise e a sua repercussão na evolução dos indicadores de audimetria, recorrendo a dados, abordagens metodológicas e a variáveis explicativas mais adequadas e/ou inéditas face ao que se verifica da literatura internacional da especialidade.

Segundo a *Figura II.3.1*, em primeiro lugar, é necessário a exposição da audiência ao *media*, variando a sua dimensão ao longo do dia e da semana, mas podendo seguir determinado padrão de evolução no decorrer do tempo. Frequentemente, a previsão deste indicador, a nível internacional, é realizada com base em análises de extrapolação dos seus valores históricos. Levanta-se a questão de qual é o seu grau de independência face à programação transmitida. A caracterização da sua variabilidade em função dos factores sazonais, ou de alterações nos factores de planeamento e no conteúdo da grelha de programação, permitirá avaliar se existem períodos onde mais facilmente esta variável possa ser prevista com base em análises univariadas. A existência de alguma estabilidade nos valores das taxas médias de audiências para o total da TV, apresenta uma grande relevância para as análises de previsão, já que poderá constituir um ponto de apoio em circunstâncias, onde se devem observar comportamentos voláteis e dinâmicos. Portanto, este deverá ser o ponto de partida para as questões a desenvolver.

Os telespectadores podem-se tornar disponíveis para ver os programas que considerem mais apelativos, em função dos conteúdos disponíveis em determinado momento. Neste contexto, a taxa de audiência para o total da televisão num dado instante, não será unicamente em função de factores externos e individuais, mas sim também, influenciada pelo conteúdo programático e da sua correspondente divulgação.

A exposição, também, poderá ser essencialmente estrutural. Previsivelmente é relevante verificar se a exposição televisiva resulta em maior grau do desenvolvimento de hábitos e rotinas denotando, uma certa passividade da audiência, ou de uma selecção mais activa na qual a estruturação da grelha e a sua divulgação poderá ser determinante.

Por vezes, são adoptadas estratégias de programação que pretendem promover a fidelização do público a determinados programas ou tipologias com o intuito de conseguir estabilizar, e maximizar, os índices de audiências no decorrer do tempo. Por exemplo, recorrendo à transmissão de séries ou telenovelas no mesmo horário ao longo de toda a semana, ano após ano, incentiva-se à lealdade programática, que também deve ser analisada. As estratégias de “*Lead_in*” ou de “*Lead_out*”, entre outras, pretendem maximizar as audiências ao longo de um maior horizonte temporal, transmitindo-se um programa menos apelativo em faixas horárias adjacentes a programas mais populares.

Em virtude da existência de efeitos de herança entre programas adjacentes, os programas mais fracos saem beneficiados, especialmente quando são transmitidos após, ou entre programas, que exercem uma maior atractividade sobre o público. Os efeitos de herança encontram-se condicionados por alguns factores, tais como, os períodos horários, a contra programação e a composição sociodemográfica da audiência. Desconhece-se a influência destes comportamentos na evolução dos *ratings*.

A análise dos fluxos de audiências entre os programas adjacentes e entre os canais é, por conseguinte, um dos factores que deve apresentar uma maior repercussão nos índices de audiências, permitindo avaliar se a variabilidade observada nos *ratings* resulta em maior grau da concorrência exercida entre os canais, ou das alterações verificadas nas taxas médias de audiências para o total da TV. A duplicação das audiências em determinados programas, tipologias e canais, também poderá condicionar a evolução de alguns índices de audiências.

Os índices de audiências traduzem exposições e, não obrigatoriamente, preferências, ou seja, apenas representam o que é seleccionado pelos telespectadores. Com o acréscimo verificado no número de opções disponíveis, as selecções podem coincidir, cada vez mais, com as preferências do público. Porém, em Portugal, a maior parte do painel não tem TV Cabo, para o período em análise. Deverá ser construído um modelo integrado e explicativo da exposição televisiva e avaliar a sua influência na evolução dos *ratings*.

A notoriedade de um canal influencia os índices de duplicação nas suas diversas vertentes, assim como as tipologias programáticas, condicionando as abordagens metodológicas mais adequadas. Também não devemos dissociar a componente comercial e a programática. A dimensão e a estruturação dos espaços publicitários, também deverá influenciar os *ratings* programáticos, uma vez que, existe uma inegável tendência para o “*zapping*”, isto é, para evitar a publicidade televisiva. A análise das exposições sociodemográficas, promovendo a identificação de comportamentos de lealdade programática, tipológica e ao canal, também pretende dar resposta às necessidades de um correcto “*targeting*” em função dos produtos divulgados nas campanhas publicitárias

Assim, atendendo aos recursos disponíveis, podem ser enumeradas as seguintes questões operacionais para o trabalho proposto e apresentado na presente tese:

(1) Identificação das metodologias e das áreas de análise estatística mais recentes que podem tornar viável uma melhor análise das bases de dados de audimetria, quer da área da previsão, quer no âmbito da classificação de dados, considerando diferentes tipologias de dados e contextos de análise. As metodologias em análise incidem na análise simbólica, no *Data Mining* e na Estatística Multivariada.

(2) Análise das taxas médias de audiências para o total da TV, uma vez que este indicador caracteriza a evolução da exposição global ao *media*.

Este índice, representa a percentagem do painel que contactou com a totalidade dos canais durante um determinado intervalo de tempo. Em função das variáveis sociodemográficas, poderá existir uma maior ou menor aproximação entre as suas rotinas de consumo televisivo e os respectivos padrões de disponibilidade e de vontade de ver TV. Pretende-se avaliar qual é o seu grau de independência face à grelha de programação, caracterizando a sua evolução interanual e para cada ano considerado nas análises. Devemos identificar os principais factores que condicionam a sua evolução.

Também se avalia a viabilidade da sua previsão, recorrendo a análises univariadas de segmentação, *clustering* e previsão dos dados. Procedemos à avaliação da influência sociodemográfica na sua evolução. Enquanto proposta para futuras análises a realizar, são sugeridas algumas abordagens metodológicas que visam a análise dos dados dos ficheiros de base de audimetria, nomeadamente, no âmbito da Regressão linear Simbólica e do *Data Mining*, assim como o cálculo de previsões de valores de *Tvr%* para suportes programáticos, em função dos valores calculados para as faixas horárias.

As taxas médias de audiências são realizadas em função dos valores de shares de audiência face ao volume global de exposições, por conseguinte, a previsão deste indicador apresenta uma grande relevância. Do ponto de vista sociodemográfico, pretendemos verificar quais são os segmentos que apresentam evoluções mais semelhantes da variável em estudo, assim como diferenciá-los ao nível da variabilidade da sua exposição televisiva, identificando os que apresentam uma maior consistência e uniformidade da exposição no decorrer da semana. Por último, avalia-se a repercussão dos espaços comerciais na sua evolução.

(3) Análise das taxas médias de audiências ou *ratings*, sendo este o indicador mais relevante do ponto de vista do planeamento de media. Pretende-se caracterizar a sua variabilidade e identificar os factores com uma maior repercussão a este nível, avaliando-se posteriormente a viabilidade da sua previsão. Apenas nos contextos em que a nível agregado for observado uma maior variabilidade, se justifica a posterior desagregação das análises.

A emissão diária deverá ser subdividida em períodos que apresentem uma evolução diferente da variável em estudo, e que devem ser objecto de análise de forma diferenciada, condicionando as tipologias de dados e metodologias mais apropriadas. Os períodos horários diferenciam-se quanto à composição sociodemográfica da audiência, estruturação e composição da grelha, estratégias de contra-programação, entre outros factores, incutindo uma evolução e variabilidade diferente nos principais indicadores. Será necessário avaliar se os índices são mais estáveis quando lidamos com emissões regulares situadas em contextos de concorrência que apresentam uma maior semelhança e, tendo em conta o conjunto global dos canais.

Também será necessário avaliar as situações em que as alterações horárias, ou os diferentes contextos de concorrência, geram nestes índices uma maior variabilidade. Determinadas tipologias programáticas podem estar mais dependentes da sua envolvente, o que exigirá instrumentos de previsão das audiências diferenciados. Em função da variabilidade observada nos dados, as abordagens de previsão diferem. Os contextos que apresentam uma maior estabilidade viabilizam as extrapolações a partir das observações históricas. Por vezes, a informação histórica utilizada em análises de previsão, encontra-se distorcida por grandes acontecimentos, como por exemplo, a transmissão de jogos de futebol. Portanto, será necessário a identificação das situações que possam produzir observações atípicas, para que estas sejam excluídas dos cálculos quando se realizam previsões para novas observações. Em simultâneo, devem ser realizadas análises de cariz exploratório aos valores de shares de audiências. Os períodos horários diferenciam-se quanto às estratégias de contra programação adoptadas, condicionando a evolução desta variável, assim como a evolução de *Tvr%*.

(4) Análise dos índices de lealdade programática, ou seja, dos índices de duplicação de audiências entre episódios ou transmissões de um mesmo programa. Estas variáveis elucidam-nos a respeito da tendência que o público apresenta para seleccionar, frequentemente, um mesmo programa ou espaço programático. Permite-nos realizar uma avaliação qualitativa da evolução dos *ratings*.

A análise da totalidade dos comportamentos de duplicação apresenta uma grande relevância, uma vez que os anunciantes precisam de conhecer estes indicadores para conseguir avaliar qual a intensidade da distribuição dos contactos.

(5) Análise dos índices de duplicação de audiências entre programas, ou entre faixas horárias, que podem ser directamente adjacentes, ou não. Simultaneamente analisar as decisões de ligar e desligar o televisor, e a análise dos fluxos de audiências entre os canais generalistas, ou perante os canais por cabo, satélite ou vídeo. Conseguir favorecer os índices de duplicação e minimizar a exposição aos canais concorrentes, permite aos canais poder apresentar em determinados períodos, uma maior estabilidade nos valores dos seus índices de audiências, facilitando a sua posterior previsão. A análise dos fluxos de audiências entre canais permite-nos avaliar qual o grau de independência do desempenho da grelha de programação de um canal face aos restantes canais, e consequentemente, avaliar a repercussão da contra programação.

(6) Para a análise de comportamentos de lealdade, uma vez que apenas dispomos de dados agregados de índices de audiências, será viável recorrer a determinados índices sociodemográficos que traduzem a exposição dos telespectadores a determinados programas, com o intuito de associar tipologias, canais e programas aos segmentos representados nas bases de dados. Por exemplo, recorre-se aos índices de adesão, atractividade, afinidade e fidelidade programática. Os índices de fidelidade, por sua vez, também podem influenciar significativamente a dimensão da duplicação de audiências entre os programas adjacentes e condicionam a quantificação dos valores de *ratings*, onde é necessário ter em atenção o grau de permanência dos telespectadores do decorrer da transmissão dos suportes incluídos nos planos de *media*.

Os comportamentos de lealdade condicionam a duplicação de audiências entre programas, influenciando a estabilidade dos principais indicadores. As análises realizadas promovem a definição do perfil de audiências de programas, tipologias e canais, assim como, do ponto de vista metodológico, visam contribuir para a segmentação das audiências televisivas.

Na literatura da especialidade não existe qualquer consenso a respeito do poder discriminante das variáveis sociodemográficas, ao nível da exposição televisiva, assim como do número de segmentos existentes e respectiva composição. Também é importante definir as abordagens metodológicas mais relevantes para a selecção dos suportes a incluir posteriormente nos planos de *media*.

(7) Análises multivariadas explicativas e de previsão de valores de *ratings*, recorrendo ao *Data Mining* e à análise Simbólica. Deverá ser avaliado quais os factores estruturais representados nas bases de audimetria com uma maior capacidade explicativa e predictiva da evolução dos *ratings*. Também devem ser comparadas as metodologias lineares, e não lineares, ao nível da precisão dos resultados proporcionados.

(8) Análise de índices de audiências comerciais que traduzem a exposição dos telespectadores às inserções publicitárias, com o intuito de desenvolver análises explicativas lineares e não lineares, que permitam a avaliação da repercussão de factores estruturais e a posterior previsão deste indicador.

O que realmente interessa aos anunciantes são os valores das audiências comerciais e a possibilidade de uma inserção em concreto, e não, de um bloco de anúncios. Será necessário reflectir sobre a diferença, ou relação existente, entre as audiências programáticas e as comerciais dos blocos de anúncios que se encontram no seu decorrer, ou entre eles. Devemos verificar se uma série cronológica de dados de audiências comerciais também apresenta a sua própria dinâmica. Em principio não terá, mas será sim, o reflexo das emissões dos programas em que se enquadram. Por conseguinte, estudar as audiências comerciais sem estudar as correspondentes audiências programáticas é inútil e, na sua sequência, será necessário avaliar se é possível prever as audiências comerciais, a partir do estudo das emissões dos programas que as enquadram.

As análises incidem nos valores dos *ratings* comerciais, uma vez que este índice determina o custo da publicidade televisiva. A relação entre as audiências comerciais e as audiências programáticas é complexa, não deverá ser linear, nem ser idêntica nos diversos momentos do dia, e dias de semana. Tem que ser avaliado o grau de linearidade da relação entre ambos, os indicadores e a influência dos factores estruturais, identificando quais as variáveis que apresentam uma maior capacidade explicativa e predictiva. Se conhecermos os factores que condicionam a evolução de um indicador, que resulta do quociente entre os valores dos *ratings* programáticos e os valores comerciais, será mais fácil tentar calcular os seus valores futuros. Se conhecermos uma série de descritores adequados, então podem-se pôr em evidência as regularidades relacionadas com estes. .

Será relevante, em especial, verificar qual a influência das tipologias programáticas, para além de outros factores situacionais. Quando se desconhece o modo de proceder dos responsáveis da programação, não é possível dizer qual será o programa que será transmitido, mas podemos tentar prever a sua tipologia. A estabilidade da tipologia de conteúdos de um canal, representa uma consequência da sua estratégia, e por conseguinte, a médio prazo podem não apresentar uma grande variabilidade.

CAPÍTULO III

FUNDAMENTAÇÃO TEÓRICA

Neste capítulo consta o estado da arte nas áreas que são mais relevantes para a construção dos modelos de exposição. Identificam-se as lacunas existentes e os principais entraves teóricos e metodológicos e estabelece-se um referencial teórico para dar suporte ao desenvolvimento do trabalho.

As referências apresentadas incidem nos modelos de “*ratings*”, uma vez que esta perspectiva de construção dos modelos de exposição deve ser privilegiada. Também se descrevem as áreas alternativas de construção dos modelos que viabilizam o investimento publicitário, estando estes últimos condicionados pelas perspectivas explicativas do comportamento dos telespectadores.

Inicialmente constam alguns capítulos que pretendem introduzir determinados conceitos que podem condicionar a compreensão dos objectivos e da estrutura das análises propostas, sendo algumas temáticas desenvolvidas nos *Apêndices*.

III.1. INTRODUÇÃO À ACTIVIDADE DE PLANEAMENTO DE *MEDIA*

A actividade de planeamento de *media* apresenta como principal objectivo optimizar o investimento publicitário, através da identificação da melhor solução estratégica e tática de aquisição de publicidade, e de acordo com os objectivos de Marketing. A TV corresponde ao *media* que é privilegiado em maior grau pelos anunciantes, sendo a audimetria essencial neste processo. Nos seguintes capítulos descreve-se em maior detalhe o processo de planificação de uma campanha publicitária.

III.1.1. As etapas que integram a construção de uma campanha publicitária

Na *Figura III.1.1.1* consta a sequência de análises que integra a construção e avaliação de uma campanha publicitária. Destacamos quatro etapas: a redacção do “briefing”, a concepção da campanha, a execução da campanha e a avaliação da sua eficácia.



Nas etapas iniciais é necessário serem realizados determinados estudos que visam o conhecimento do mercado onde se insere a marca e do perfil de consumo de *media* do segmento-alvo, para além da análise do investimento publicitário realizado por parte das empresas concorrentes. Os estudos prévios permitem que sejam definidos os objectivos de *media* ao nível do horizonte temporal e âmbito geográfico da campanha, assim como os valores mais adequados para os indicadores que condicionam a aquisição de espaço comercial.

Fig. III.1.1.1 – As etapas que integram a construção de uma campanha publicitária.
Fonte: Elaborado pela autora.

No *briefing* deverá constar a análise do produto, concorrência e do consumidor, assim como a descrição das estratégias de Marketing do anunciante e os objectivos e limitações da publicidade. Ao nível da concepção da campanha deve ser realizada a criação publicitária, a escolha dos canais de comunicação, quer dizer, a selecção dos *media* e dos suportes (plano de *media*) e a negociação e compra do espaço e tempo publicitário.

Neste processo existe a necessidade de investigar os *media* e os suportes publicitários, ao nível do perfil e dimensão das suas audiências. O que se denomina de investigação ou estudos de audiência, num contexto de planeamento de *media*, engloba especificamente o estudo de questões relacionadas com a exposição aos meios de comunicação.

Quando se opta pela TV como veículo publicitário, os modelos de exposição, optimização e de avaliação dos efeitos da publicidade nas vendas e ao nível da divulgação da marca permitem que seja proposto um plano de investimento em publicidade televisiva e designam-se de modelos de *media*. Quem intervém no planeamento de uma campanha deverá compreender os padrões de consumo televisivo e a evolução dos indicadores que deles resultam, para além de tentar prever a exposição do segmento-alvo considerando em maior grau o curto prazo.

O processo conclui com a execução, acompanhamento e a avaliação dos resultados por parte do anunciante que apenas será possível recorrendo a determinados indicadores de audimetria. As estações televisivas em determinados contextos devem compensar os anunciantes quando a exposição efectivamente alcançada é inferior ao que foi previsto e segundo a qual se definiu a estrutura de preços para o plano de investimento anteriormente adquirido. As campanhas são essencialmente avaliada em função dos valores das audiências acumuladas e dos *ratings* que conseguiram alcançar.

III.1.2. A implementação do planeamento de *Media*

Quando se escolhe um canal de comunicação, tem que ser considerada a cobertura do alvo, a repetição da mensagem e a adequação da mensagem ao canal seleccionado. A TV apresenta múltiplas vantagens, face aos restantes meios de comunicação, quando se pretende comunicar com o segmento-alvo de uma campanha publicitária. Destacamos o seu elevado alcance, um reduzido custo por contacto e uma maior qualidade da medição das variáveis publicitárias.

O seu carácter audiovisual confere-lhe grandes vantagens, reforçando a notoriedade de uma marca e influenciando o comportamento dos consumidores. Por conseguinte, a TV constitui um dos principais veículos publicitários e onde, em geral, incide o maior investimento realizado pelos anunciantes.

Alguns dos critérios que condicionam a selecção dos *media* e suportes publicitários correspondem aos seguintes: ao grau de penetração e perfil da sua audiência, ao grau de duplicação das audiências entre os suportes, aos níveis de saturação publicitária, ao tempo médio de contacto de um indivíduo face ao suporte, aos custos da inserção e à flexibilidade na sua contratação. Também são consideradas as características técnicas do *media*, a afinidade do público-alvo da campanha face ao *media* em avaliação e as eventuais sinergias entre os diversos *media* seleccionados. Entre os critérios qualitativos destacamos: a credibilidade, prestígio e capacidade da influência do *media* e do suporte, a sua imagem e posicionamento, a forma de recepção da mensagem e ser um meio gratuito ou pago.

A estratégia de *media* tem início com a escolha do canal, formatos e sazonalidade da campanha. Posteriormente, inicia-se a preparação do plano de *media*, isto é, a selecção dos suportes, como, por exemplo, dos canais ou períodos horários, a definição dos locais onde os adquirir e do calendário exacto das inserções. Os modelos de *media*, em especial os modelos de exposição e de optimização, servem de suporte à construção dos calendários de investimento em publicidade. Conclui-se com a compra do espaço e com o controlo da campanha, onde se verifica se as inserções estão de acordo com o que consta do plano e se os seus níveis de audiência correspondem ao que foi negociado.

Na área televisiva a aquisição dos suportes publicitários pode ser realizada em diferentes momentos e mercados que se diferenciam ao nível dos custos e garantias proporcionadas por parte das estações televisivas. Numa fase inicial designada de “up-front market” podem ser definidos planos de investimento que integram um conjunto estruturado de suportes para várias semanas e que resultam da aplicação de determinados modelos de *media* de apoio à decisão.

Nestes mercados são transacionadas as inserções dos períodos de maiores audiências e os anunciantes são compensados quando não são alcançados os níveis de exposição previstos. Também é possível adquirir um menor conjunto de suportes no mercado das inserções, em função das unidades ainda disponíveis e a preços mais reduzidos ou no “scatter market” onde a publicidade poderá ser adquirida através da construção de planos de *media*, mas com menores custos face aos que são praticados no “up-front market”. Em função do mercado escolhido, os indicadores que definem em maior grau o custo da publicidade podem-se diferenciar.

III.1.3. OS MODELOS DE *MEDIA*: OS MODELOS DE EXPOSIÇÃO, OPTIMIZAÇÃO E DE AVALIAÇÃO DOS EFEITOS DA PUBLICIDADE

Um plano de investimento em publicidade televisiva tem início com a escolha do público-alvo, o que poderá implicar a segmentação do painel e uma análise do seu perfil de consumo de *media*.

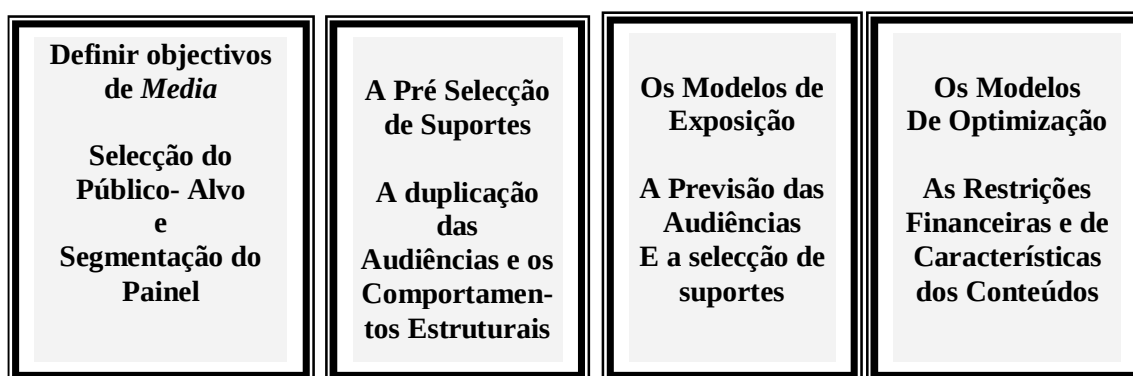


Fig. III.1.3.1 – Etapas que integram a construção de um plano de investimento em publicidade televisiva.

Fonte: Elaborado pela autora.

Posteriormente, devemos proceder a uma pré-selecção ou redução da dimensionalidade do número de suportes em função da associação dos conteúdos transmitidos aos segmentos representados nas bases de dados de audimetria e dos resultados das análises realizadas às audiências conjuntas que permitem avaliar os comportamentos estruturais de consumo televisivo.

Os modelos de exposição, isto é, de previsão das audiências devem incidir na subamostra de suportes previamente identificada. As restrições orçamentais, entre outras que resultem do conhecimento existente sobre o perfil de consumo de *media* do segmento-alvo, devem ser consideradas nos algoritmos de optimização (*Fig. III.1.3.1*).

Inicialmente, utilizavam-se os inquéritos como instrumento de previsão. Após o desenvolvimento da audimetria, as análises começaram a incidir nas bases de dados que caracterizam a exposição do painel no decorrer do último mês, pressupondo a estabilidade dos comportamentos de consumo televisivo, e neste contexto recorrendo apenas à optimização. Uma vez que o sector tem evoluído e o telespectador dispõe de uma cada vez maior oferta televisiva e de alternativas de ocupação de tempos “vazios” e de lazer poderá não se verificar essa estabilidade e novas perspectivas e modelos matemáticos são necessários recorrendo às bases de audimetria. Neste contexto os modelos de exposição são essenciais para o investimento em publicidade televisiva.

Os modelos explicativos da interacção do painel com o meio televisivo ou dos índices de audiências podem contribuir para as análises de previsão a diversos níveis. Quando as variáveis identificadas como principais condicionantes da evolução da variável critério forem quantificáveis, conhecidas ou calculadas *a priori* com alguma precisão, podem ser consideradas em análises de previsão. Caso contrário, demonstra-se a imprevisibilidade dos indicadores em estudo e algumas das suas causas e podem ser avaliadas perspectivas de análise alternativas. Quanto melhor se entender a complexidade de um sistema que apresente componentes inter-relacionados melhor será a decisão e, neste contexto a selecção das perspectivas de construção dos modelos de exposição.

Nesta área, nunca se poderá definir como objectivo a realização de previsões deterministas, mas, sim, a identificação e análise das dimensões que condicionam maioritariamente a exposição e a evolução dos indicadores que definem o custo da publicidade televisiva.

Sómente após a avaliação da influência de determinantes estruturais e individuais na exposição televisiva e na evolução dos principais indicadores de audimetria se poderá definir uma sequência metodológica que permita contribuir para a temática da previsão das audiências, e que será função dos contextos que estiverem em análise. Na análise dos comportamentos de consumo televisivo é relevante compreender quais são os factores que determinam a exposição, uma vez que os resultados condicionam a estruturação da grelha e os conteúdos a incluir nos planos de *media*.

Os indicadores acumulados de “*ratings*” permitem caracterizar a exposição televisiva, nomeadamente a sua regularidade que é indispensável para que sejam viáveis as análises de indução. Também, permitem a interpretação da evolução dos indicadores agregados de audimetria e a selecção dos suportes a incluir em planos de *media*. Os modelos de exposição explicativos e de previsão de valores de *ratings* programáticos e comerciais ou da exposição televisiva agregada corresponde a uma área essencial para a gestão e sobrevivência das empresas ligadas ou dependentes do sector televisivo.

Em suma, na actividade de planeamento de *media* são necessários três tipologias de modelos: os modelos de exposição, os modelos de optimização e os modelos de avaliação dos efeitos da publicidade.

- Os **modelos de exposição** têm como função compreender e prever a exposição televisiva.

Podem incidir apenas na perspectiva de quantificação da cobertura e frequência de um plano ou serem mais ambiciosos. A cobertura e a frequência permite o cálculo das audiências acumuladas, ou seja, do número de telespectadores expostos aos suportes após sucessivas transmissões dos mesmos. Estes modelos devem permitir a avaliação do grau de duplicação das audiências entre os suportes, uma vez que a mensagem deve ser repetida ao alvo para ser eficaz e incentivar a aquisição dos produtos e serviços divulgados. Uma maior percepção sobre o perfil de consumo de *media* do segmento alvo é essencial para a sua construção.

- Após o desenvolvimento dos modelos de previsão das audiências, os planos de *media* são construídos através da aplicação de **algoritmos de optimização**.

Estes modelos pretendem seleccionar o melhor conjunto de suportes, maximizando a exposição ao plano delineado, mas considerando determinadas restrições que podem ser económicas, ou ao nível da frequência alcançada, ou representar determinadas características dos suportes. Por vezes, os suportes são seleccionados segundo o conceito da “frequência efectiva” (*Murray & Jenkins, 1992*) que é contestado por inúmeros autores que defendem, sim, a necessidade de recorrer à “frequência óptima” (*Cannon & Riordan, 1994*) ou frequência natural e característica de um segmento (*Zarate, 1997*). Todavia, alguma bibliografia considera que apenas uma exposição poderá ser suficiente e, por conseguinte, será mais relevante a cobertura do público-alvo (*Longman, 1997*).

Os modelos de exposição contribuem para uma melhor definição das restrições que devem ser impostas a estes algoritmos, para além, de condicionar outras características que os modelos de optimização assumem em etapas posteriores das análises que viabilizam a construção de um plano de investimento em publicidade televisiva.

Algumas publicações fazem incidir os algoritmos de optimização directamente na base de dados que caracteriza a exposição do mês anterior (*Craig & Ghosh, 1993*) e o utilizador deverá definir as restrições ao nível de custos, canais, tipologias e faixas horárias. Neste contexto assume-se a estabilidade dos comportamentos e não é realizada qualquer análise de previsão ou de selecção dos suportes mais adequados em função da análise do perfil de consumo televisivo do mercado-alvo.

Em *Monkman (1996)* refere-se a existência de sistemas informáticos que seleccionam aleatoriamente alguns planos de *media* que cumprem as especificações que foram definidas pelo utilizador e estimam os valores da cobertura e frequência que podem ser alcançados, assumindo que são idênticos ao que se verificou no histórico.

Por vezes, poderá ser atribuída uma probabilidade de exposição para cada dia/canal/periodo e que são utilizadas para calcular o valor da cobertura de um plano que satisfaz as restrições definidas ao nível dos custos. Estes modelos apresentam inúmeras limitações, entre as quais, o facto das previsões a nível individual dificilmente conduzirem a resultados com alguma precisão.

A solução proporcionada poderá não ser a mais adequada, uma vez que não é realizada qualquer análise de identificação dos suportes mais adequados e o valor final proporcionado corresponde à média da totalidade dos planos que satisfazem as restrições, não sendo, consequentemente, representativa da variabilidade dos resultados que podem conduzir a esse valor. Estes modelos também apresentam outras limitações, como, por exemplo, não contribuírem para a definição do valor da frequência. Segundo *Makienko (2012, p.61)*, cada marca deve encontrar o seu valor de frequência, sendo em geral este o aspecto que é privilegiado, uma vez que os anunciantes em geral dispõe de reduzidos orçamentos: “*Media plans with low budgets seeking maximum reach won’t be able to overcome the threshold frequency level for the advertising to become effective.*”

➤ Por último os **modelos sobre os efeitos da publicidade** pretendem avaliar em que medida a exposição às campanhas publicitárias consegue modificar os comportamentos dos telespectadores e incentivar a aquisição dos produtos e serviços divulgados. Avaliam qual é a sua repercussão no volume de vendas ou no número de serviços contratados e nas atitudes dos consumidores.

III.1.4. O plano de *Media*

A construção de um plano de *media* implica a selecção dos suportes e a sua calendarização, apoiando-se nas bases de dados que caracterizam as audiências e em determinados algoritmos que viabilizam a selecção do conjunto de inserções que podem ser mais eficazes para determinada campanha.

Os dados dos ficheiros base de audimetria (“*raw data*”) viabilizam a construção de duas bases de dados de índices de audiências, as quais podem ser consultadas recorrendo aos *softwares* Telereport e *Publitem*. O *software* Telereport permite aceder aos dados das audiências programáticas e para blocos de anúncios, enquanto que o sistema *Publitem* viabiliza a análise das inserções transmitidas nos intervalos publicitários.

III.1.4.1. Os indicadores de desempenho do plano

Para entender melhor a estrutura de um plano de *media* será necessário definir adequadamente a terminologia utilizada sectorialmente. A seguinte *Figura III.1.4.1* resume alguns dos indicadores mais relevantes, os quais podem ser classificados em medidas brutas ou agregadas e em medidas acumuladas. Na primeira classificação podem ser incluídos os *ratings* ou taxas médias de audiências, os *shares* e o custo por mil contactos. De entre as medidas acumuladas destaca-se o *cume*, ou seja, o número de pessoas que contactam com um suporte pelo menos cinco minutos num período de tempo e a frequência que expressa o número de vezes que alguém esteve exposto ao *media*.

Os principais indicadores de audimetria, e dos quais depende fortemente o investimento realizado em publicidade televisiva, são as taxas médias de audiências (*Rat%*), as taxas médias de audiências para o total da televisão (*Tvr%*) e os *shares* de audiência (*Shr%*).

O *rating* (*Rat%*) representa a audiência de um evento como uma percentagem sobre a base do público-objectivo. Se a audiência de um programa foi de 30%, isto significa que sendo um valor médio, em teoria 30% da população considerada assistiu a esse programa. Este índice representa a percentagem de pessoas que viram um evento sobre o total de pessoas com possibilidade de ver TV e consideradas como público-objectivo, isto é, representa a audiência útil entre o público objectivo, ou seja, expressa a percentagem de audiência útil de um evento que foi alcançado com uma possível inserção nesse evento.

TAXA MÉDIA DE AUDIÊNCIA OU <i>RATING</i> Cada indivíduo é ponderado relativamente ao tempo de contacto com o programa <i>RAT%</i> (evento) = $\frac{\text{Total de Seg. Contactados nesse evento}}{\text{Duração do período}}$	COBERTURA DE UM PLANO Exemplo; 25% do alvo não contactou e 75% contactou com a campanha, sendo essa a cobertura do plano.
TAXA MÉDIA DE AUDIÊNCIA PARA O TOTAL DA TV OU <i>TVR%</i> N.º de indivíduos que contactaram com todos os canais durante a transmissão de determinado evento. $TVR\% = \left(\frac{\text{Total de Segundos contactados no total de televisão}}{(\text{Duração do evento} * \text{n.º de indivíduos do alvo})} \right) * 100$	FREQUÊNCIA MÉDIA DE CONTACTOS OU <i>OTS</i>: NÚMERO MÉDIO DE CONTACTOS Exemplo: <i>OTS</i> = 2,5. Em média o alvo contactou duas vezes e meia com a campanha publicitária.
<i>SHARE</i> DE AUDIÊNCIA OU <i>SHR%</i> Quota de audiência de um evento Face ao global de exposições ao <i>Media</i> $SHR\% (\text{evento}) = \frac{RAT\% (\text{Evento})}{TVR\% (\text{Evento})}$	GROSS RATINGS POINT OU <i>GRP</i> O <i>GRP</i> é equivalente ao produto da cobertura pelo n.º médio de contactos conseguido pela campanha ou ao somatório dos <i>ratings</i> de cada uma das inserções que integram o plano.

Figura III.1.4.1 – Caracterização de alguns indicadores de audimetria.
Fonte: Elaborado pela autora.

A relevância dos *ratings* advém de viabilizar a quantificação dos valores das audiências acumuladas representadas pelo índice GRP e de definir o custo da publicidade no mercado das inserções. Quando se seleccionam os suportes publicitários será conveniente seleccionar os que apresentam maiores valores de *rating*, mas tendo em atenção que grande parte da audiência não deverá coincidir entre eles, caso contrário contribuirá para um acréscimo do valor da pressão publicitária.

A taxa média de audiência para o total da TV (*Tvr%*) designa a percentagem do painel que tem o televisor ligado em determinado momento sobre o público objectivo considerado. O share de um evento (*Shr%*) representa a sua audiência como uma percentagem do número de pessoas que estão a ver TV em determinado momento. O valor do *share* será a percentagem de participação de um evento sobre o total da audiência dos canais, representando um maior ou menor número de indivíduos em função da audiência total do *media* que varia em função do dia e hora sobre a qual se realiza o cálculo.

No planeamento de *media* também são relevantes os valores de GRP, a frequência média ou “oportunity to see” (OTS), a cobertura e a distribuição de contactos. A cobertura designa a percentagem do publico-alvo que contactou com a campanha pelo menos uma vez, independentemente do número de contactos. A distribuição de frequências representa os diversos graus de exposição do público-objectivo à campanha publicitária. A frequência média corresponde ao número médio de vezes que os indivíduos contactados do nosso segmento-alvo tiveram a oportunidade de ver uma inserção e resulta do quociente entre o valor de GRP e a cobertura. A frequência efectiva faz referência ao número médio de vezes considerado como necessário para provocar uma resposta positiva no segmento-alvo da campanha publicitária. O indicador GRP que relaciona a cobertura com o valor da frequência corresponde à soma bruta de audiências de uma campanha conseguidas durante determinado período de tempo para o público-alvo. Este índice resulta do produto da cobertura pelo número médio de contactos ou do produto da audiência média da campanha pelo número de inserções (*Farris & Parry, 1991*).

O GRP, também, pode ser definido como a soma dos *ratings* individuais de uma combinação de suportes em determinado período de tempo (*Dickson, 1991*) ou como o total de impactos produzidos por uma campanha sobre o nosso público objectivo. Representa uma medida de pressão, uma vez que resulta da combinação do alcance, cobertura e frequência alcançados pela campanha. Não têm em conta o número de duplicações e de exposições repetidas e os indivíduos contabilizam-se tantas vezes quantas são expostos ao anúncio ou inserção.

O conceito de cobertura tem em conta as duplicações de audiências, triplicações, etc. As duplicações correspondem a repetições de audiências entre suportes e depende dos diferentes perfis que constituem o universo televisivo, mas facilmente se percebe que será lógico ver vários programas e canais em diferentes horários e sessões televisivas. As implicações destes comportamentos são cruciais para o planeamento de *media*. Quanto maior for o número de suportes combinados, com uma audiência que não seja coincidente, mais fácil será contactar com o total do público-objectivo pelo menos em uma ocasião. Um planificador tem que alcançar os níveis de cobertura definidos, mas por vezes os suportes seleccionados apresentam uma elevada audiência partilhada.

A audiência bruta representa o número de indivíduos que contactaram com uma combinação de suportes pelo menos uma vez, tendo em conta as duplicações que se produzem entre os diversos suportes. A audiência bruta acumulada designa o número de pessoas que viram um anúncio pelo menos uma vez numa combinação de suportes, tendo em conta as duplicações entre suportes diferentes e as diferentes inserções num mesmo suporte, ou seja, a acumulação da audiência.

Nas bases de audimetria também constam índices que permitem avaliar o perfil das audiências dos suportes e a sua “lealdade”. Por exemplo, o coeficiente de afinidade indica se um evento é idóneo para determinado público-objectivo, uma vez que é calculado como a percentagem da audiência total do suporte que representa a audiência do público-objectivo.

O coeficiente de afinidade permite conhecer em que medida o nosso alvo assiste ao suporte comparado com a população total, indicando que suportes apresentam um menor desperdício de contactos relativamente ao público-objectivo da campanha.

A audiência útil corresponde ao número de pessoas que sendo audiência de um suporte pertencem também ao público-objectivo, uma vez que apresentam as características que foram definidas para o alvo da campanha. A ponderação de um alvo denota o interesse que determinados segmentos apresentam para o anunciante ou agência de meios.

III.1.4.2. Métodos de construção do planos e sua avaliação

Uma vez seleccionados os suportes, tem que ser estabelecida a distribuição das inserções. O investimento poderá ser realizado de forma uniforme em vários períodos quando se pretende desenvolver uma campanha de manutenção de uma marca. Em alternativa podem ser intercalados períodos de inactividade com períodos de investimento uniforme, crescente ou decrescente. Os períodos de inactividade podem ser substituídos por períodos de menor investimento e uniformes no decorrer do tempo. Em campanhas de lançamento de produtos o investimento pode ser realizado em sucessivos períodos de forma decrescente e de forma inversa quando se pretende criar nos consumidores alguma expectativa face ao produto.

O planificador de uma campanha constroi vários planos alternativos e simula o seu desempenho, com a ajuda de determinados sistemas informáticos. Os planos são essencialmente avaliados ao nível da cobertura, número médio de contactos ou OTS e através da distribuição estatística dos OTS sobre o público-alvo. Os suportes, também, podem ser avaliados em função da sua audiência útil, custo por contacto útil e da sua afinidade face ao alvo a que se direccionam.

III.2. OS MODELOS DE EXPOSIÇÃO NO PLANEAMENTO DE *MEDIA*

III.2.1. Os modelos que antecederam o desenvolvimento da audimetria

Os modelos de exposição inicialmente desenvolvidos tinham como objectivo conseguir prever se determinado telespectador contacta com um suporte e a sua frequência no decorrer do horizonte temporal em que se desenrola uma campanha publicitária. A perspectiva que antecedeu o desenvolvimento da audimetria consiste em a partir dos valores das frequências de contacto individuais ser calculada a distribuição de frequências para o segmento-alvo perante o conjunto de suportes seleccionados.

A quantificação e previsão das audiências era realizada através de inquéritos diários no decorrer de uma semana e a partir do qual se definiam as frequências de contacto ou por observação ou recorrendo ao valor de frequência declarada pelo telespectador. Uma descrição mais exaustiva e uma crítica aos métodos pré-audimetria pode ser consultada em *Zarate (1997)*. Segundo esta autora, qualquer uma destas abordagens parte de pressupostos que não são válidos, ou seja, a existência de regularidade da exposição e da interpretação frequentista da probabilidade que na realidade não se ajusta aos comportamentos sociais. Também refere que não podemos calcular a probabilidade de uma acção por análise de alguns comportamentos isolados. O processo conclui, recorrendo a algoritmos de simulação que viabilizam a construção da distribuição de frequências de contacto face ao plano de *media*. Esta a perspectiva de análise imperou na bibliografia da especialidade durante décadas.

Estas abordagens apresentam inúmeras limitações, entre as quais a inviabilidade de avaliação da duplicação das audiências Os suportes devem ser seleccionados de modo a apresentarem audiências conjuntas, uma vez que a mensagem deve ser repetida ao público-alvo para ser mais eficaz e devem apresentar a maior fidelidade e lealdade por parte do seu público. Porém, através dos instrumentos existentes antes do desenvolvimento da audimetria não era possível conhecer esta informação.

III.2.2. O contributo da audimetria para a construção dos modelos de exposição

A audimetria viabilizou uma medição mais rápida, detalhada, contínua e fiável da exposição televisiva, disponibilizando ao sector um vasto conjunto de indicadores. A audimetria permitiu demonstrar, internacionalmente, que pode existir irregularidade no consumo televisivo uma vez que permite realizar uma análise diacrónica da audiência e caracterizar a exposição ao segundo. Por conseguinte, actualmente, representa o principal instrumento de previsão e de planificação.

Inicialmente, recorrendo a fórmulas *ad hoc* de duplicação e a modelos multinomiais de Dirichlet, procedia-se ao cálculo da distribuição de frequências de contacto. Alguns autores, também fazem incidir os algoritmos de optimização na base de dados de audimetria que caracteriza a exposição do mês anterior, sem previamente realizar qualquer análise de previsão (*Craig & Ghosh, 1993*). Contudo, uma vez mais, também, não é correcto assumir que a exposição no curto prazo será idêntica à histórica, ainda que os dados permitam uma caracterização mais detalhada e objectiva da exposição.

Atendendo às limitações das anteriores abordagens tiveram que ser identificadas perspectivas alternativas. Os planos de *media* podem ser construídos recorrendo à previsão de valores de *ratings*, sendo esta a perspectiva privilegiada pelo sector em alternativa às abordagens qualitativas que incidem, essencialmente, nos contextos de maior irregularidade dos dados. As abordagens qualitativas incluem as sondagens de intenção de ver e os painéis de opinião, apresentando inúmeras limitações.

O somatório dos valores de *ratings* de cada suporte considerado num plano representa um indicador que quantifica as audiências acumuladas e que também pode ser calculado como o produto da cobertura e frequência de exposição de um plano.

O conceito de frequência tem evoluído no decorrer do tempo, sendo inicialmente adoptada a frequência média, seguindo-se a frequência efectiva, frequência óptima e por último a frequência estrutural e característica de determinado segmento. A frequência efectiva representa uma exposição no mínimo de três, mas, posteriormente, foi demonstrado que o seu valor deve depender fortemente das características de uma campanha publicitária, nomeadamente da lealdade à marca, perfil sociodemográfico do segmento, complexidade da peça creativa, entre muitas outras variáveis. Não existe um único valor e transversal a qualquer campanha publicitária que contribua para uma maior eficácia do plano de investimento e a frequência necessária depende dos níveis de atenção que lhe são dirigidos. Os valores de *ratings* também condicionam a venda dos suportes publicitários no mercado das inserções.

AQUISIÇÃO DE SUPORTES EM PLANOS	AQUISIÇÃO DE SUPORTES NO MERCADO DAS INSERÇÕES
PREVISÃO DE <i>RATINGS</i> É A PERSPECTIVA PRIVILEGIADA PELO SECTOR $RAT = TVR \times SHR$ E $\sum RAT = GRP$	PREVER VALORES DE <i>RATINGS</i> OU AS AUDIÊNCIAS BRUTAS OU ACUMULADAS
OU	
COBERTURA E A FREQUÊNCIA ALCANÇADA RECORRENDO AO HISTÓRICO OU A MATRIZ DE EXPOSIÇÃO FUTURA INDIVIDUAL $GRP = COB. \times FREQ.$	SUPORTES SELECIONADOS EM FUNÇÃO DA DIMENSÃO DA DUPLICAÇÃO DAS AUDIÊNCIAS E DOS COMPORTAMENTOS ESTRUTURAIS DE CONSUMO TELEVISIVO
OU	
IDENTIFICAR PERSPECTIVAS ALTERNATIVAS DE IDENTIFICAÇÃO DOS SUPORTES PUBLICITÁRIOS MAIS ADEQUADOS A UM SEGMENTO-ALVO	A QUANTIFICAÇÃO DA DUPLICAÇÃO DAS AUDIÊNCIAS PODE CONTRIBUIR PARA PREVER AUDIÊNCIAS BRUTAS E ACUMULADAS

Fig. III.2.2.1 - Os indicadores que condicionam o investimento em publicidade televisiva.

Fonte: Elaborado pela autora.

A *Figura III.2.2.1* representa as necessidades de informação do sector no âmbito da construção dos planos de investimento em publicidade.

A previsão de valores de *ratings* pode ser realizada através de análises univariadas ou multivariadas, mas que no geral também têm apresentado resultados insatisfatórios. Dificilmente, será viável a construção de uma base de dados de exposição “determinista” para o curto prazo, considerando as exposições a nível individual que permita a previsão da cobertura e da distribuição de frequências. Os comportamentos a nível individual são classificados como acausais, indeterminados e imprevisíveis. Consequentemente, é essencial e devem ser desenvolvidos esforços no sentido de serem alcançados os melhores resultados no âmbito dos modelos explicativos e de previsão de valores de “*ratings*”. Verificamos que são muito escassos os modelos quantitativos alternativos aos modelos de *ratings*, por exemplo, através da quantificação da exposição de um segmento face a um conjunto de suportes e que por sua vez dependerá das perspectivas teóricas explicativas da exposição.

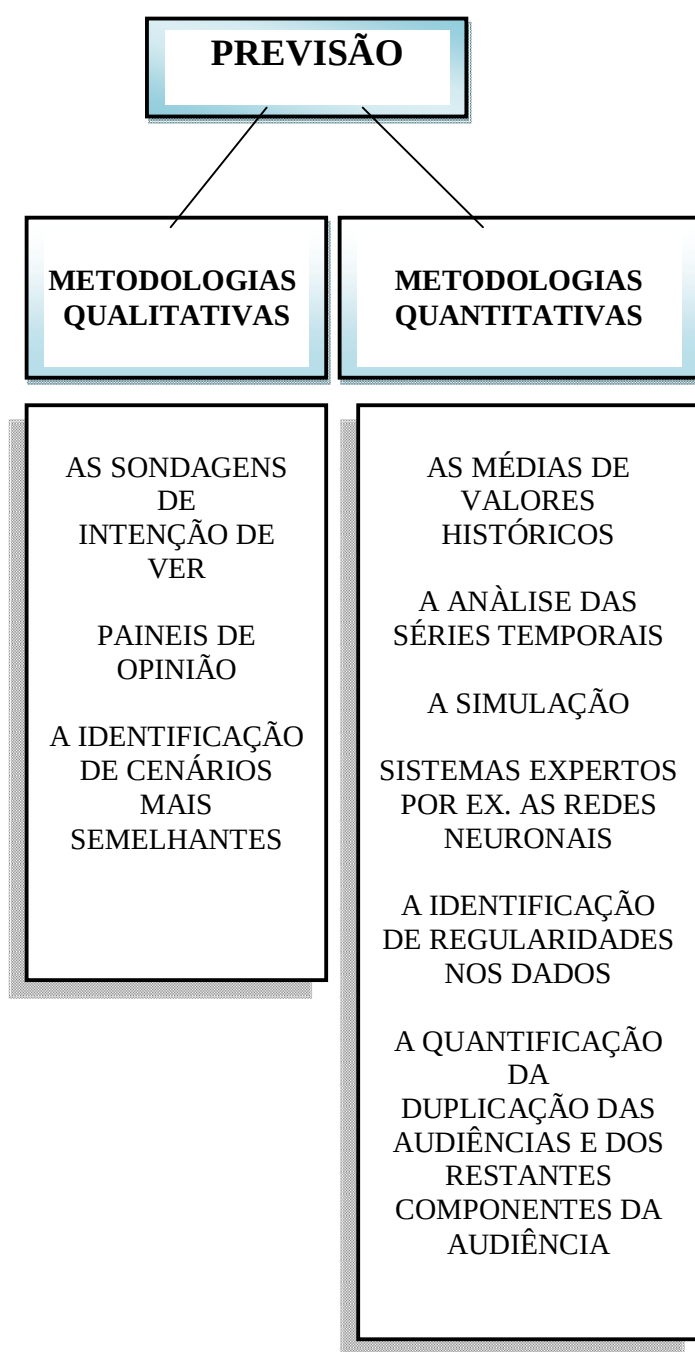
Em alternativa aos modelos de *ratings* seria necessário a quantificação das audiências totais, acumuladas ou brutas do conjunto de suportes que devem apresentar alguma duplicação das audiências ou a identificação de perspectivas de análise que viabilizem uma selecção mais efectiva dos suportes por interpretação da estrutura dos comportamentos de consumo televisivo.

III.2.3. As metodologias utilizadas na construção dos modelos de exposição: Vantagens e limitações

As áreas metodológicas que viabilizam a construção dos modelos de exposição podem ser diferenciadas em análises qualitativas e quantitativas. As análises quantitativas assentam na análise dos dados históricos sobre o sistema em avaliação e projectam no futuro esses comportamentos. Por conseguinte, em geral, partem do pressuposto de que irá prevalecer no futuro as condições que também determinaram no passado a sua evolução.

As análises qualitativas são utilizadas para as previsões de longo prazo ou quando não existe estabilidade nas condições subjacente e, por conseguinte, são menos objectivas. Na área da audimetria (*Fig. III.2.3.1*) incluem as sondagens de intenção de ver e os painéis de opinião.

A audimetria viabilizou a construção de modelos matemáticos causais e não causais que



podem ser classificados nas seguintes categorias: análise das séries temporais dos indicadores de audimetria, modelos de simulação, extracção de regularidades estatísticas (estatística multivariada), os sistemas expertos tais como as Redes Neurais, as análises multivariadas lineares ou não lineares que consideram factores explicativos da evolução dos dados conhecidos ou previsíveis e a construção de modelos que permitam considerar a evolução dos fluxos de audiências e a duplicação das audiências entre os suportes publicitários. Os métodos qualitativos utilizam-se quando os dados históricos não nos proporcionam uma informação directa sobre o fenómeno que se encontra avaliação.

Fig. III.2.3.1 - As metodologias que permitem a construção dos modelos de exposição.
 Fonte: Elaborado pela autora.

Por conseguinte, as análises qualitativas são mais úteis quando os dados são escassos, ou quando pela incerteza e irregularidade dos dados e do ambiente, a extrapolação assume uma validade muito duvidosa. Todavia, na área da audimetria não podemos restringir as abordagens de previsão às perspectivas qualitativas de análise dos dados.

O método da analogia histórica pode ser utilizado quando é conhecida a evolução que decorre da introdução de produtos ou serviços análogos e tem sido utilizado nesta área. Porém, actualmente poderá não ser a perspectiva mais adequada, uma vez que a exposição é, frequentemente, não intencional, irregular e pode ser independente dos conteúdos. Apenas se os dados agregados apresentarem uma evolução mais estável, poderá ser considerada e, por conseguinte, invalida a sua aplicação nos períodos onde a exposição é superior, mas também mais variável.

Nesta abordagem compara-se a situação dos programas a prever com outras situação que possam ser consideradas semelhantes e que se verificaram no passado. Os factores a considerar abrangem, por exemplo, o tipo de conteúdo, o canal, variáveis temporais (dia, hora, mês), características da programação em contra-programação e relacionadas como o perfil sóciodemográfico do público. Em função das situações mais semelhantes observadas no passado atribui-se como previsão da audiência do programa a média observada para o subconjunto de programas previamente identificados. Na análise de Regressão linear também se prevê o valor mais provável de uma variável, ou seja, o valor em torno do qual é mais provável que se encontre o seu valor real e corresponde a um valor médio que poderá ser associado a todas as observações que apresentam equivalentes valores das variáveis explicativas incluídas no modelo.

A nível internacional esta abordagem é pouco utilizada devido à dificuldade em encontrar no passado situações muito semelhantes à actual, sobretudo se quisermos considerar muitas características da grelha de programação, entre outras variáveis. A nível nacional a grelha nos períodos de maiores audiências, por vezes, apresenta alguma semelhança ao nível das tipologias transmitidas, porém a irregularidade da exposição ao *media* poderá inviabilizar estas análises, assim como as alterações horárias de transmissão dos programas.

Considerando as abordagens estatísticas mais tradicionais e se nos concentrarmos apenas nas observações mais recentes, onde eventualmente será mais fácil encontrarmos situações mais semelhantes à actual, poderá não existir um número significativo de casos para se conseguir alcançar alguma precisão ao nível dos resultados das análises de previsão. Por outro lado, na informação histórica de longo prazo, será mais difícil considerar situações equivalentes. O passado deverá ser estruturalmente diferente do momento actual devido às contínuas alterações que se observam no cenário dos *media*: o lançamento de novos programas ou tipologias, as situações de concorrência entre canais, a modificação da constituição do painel, etc. Consequentemente, neste âmbito as análises devem considerar dados *fuzzy* ou simbólicos que permitam ter em atenção a variabilidade observada ao nível dos factores que influenciam em maior grau a evolução das audiências.

As desvantagens desta perspectiva em princípio serão ainda superiores com as alterações que se perspectivam para o sector audiovisual. Quando existir um crescente número de telespectadores com acesso a uma grande diversidade de canais será mais difícil encontrar situações consideravelmente semelhantes, a não ser que se aumente a dimensão do painel com todos os custos que lhe estão associados. A sua utilidade será, talvez, maior para tentar prever indicadores em situações gerais que apresentem poucas características, nas quais não exista a influência de factores que façam com que a situação actual em nada se pareça ao que se verificou no passado. Nesta abordagem, a inferência traduz-se na procura de uma analogia substântiva.

Os sistemas qualitativos das sondagens de intenção de ver, frequentemente, também apresentam grandes margens de erro pelos mesmos motivos. A exposição poderá não estar associada a uma intencionalidade e preferência e quando está poderá não ser conhecida e previsível com a necessária antecedência por parte de cada telespectador e de quem constrói as campanhas publicitárias. Nas estações televisivas verifica-se, por vezes, uma tendência para ocultar a programação, para além dos problemas logísticos associados a este processo e a falta de capitalização do conhecimento que o convertem num dos métodos mais onerosos.

Ainda que nos períodos que apresentam maiores níveis de audiência, as grelhas dos canais privados generalistas sejam conhecidas durante o período de transmissão das telenovelas, a irregularidade da exposição ao meio televisivo e os restantes factores estruturais podem inviabilizar que bons resultados possam ser alcançados por esta via. Este método tem sido, essencialmente, utilizado nos períodos de maiores audiências e onde os dados são mais irregulares, mas não produzem bons resultados, uma vez que os telespectadores acabam por manifestar em maior grau os seus interesses e não têm em atenção os inúmeros factores que influenciam o consumo televisivo e que são, frequentemente, desconhecidos pelos mesmos. Quando se pretende calcular as audiências totais de um conjunto de suportes através da quantificação das audiências por eles partilhada e do número de telespectadores não duplicados encontramos as mesmas limitações.

As previsões qualitativas recorrendo à intuição e a juízos de valor de painéis de opinião também devem apenas complementar as análises e apresentam uma maior relevância quando se pretende a análise de novas grelhas de programação e a sua posterior gestão e em menor grau para a actividade de planeamento de *media*, onde é necessário calcular uma grande diversidade de indicadores e recorrer à análise de bases de dados de elevada dimensão.

As abordagens quantitativas de previsão são inúmeras, desde as mais simples médias para valores históricos e extrapolações ou prolongações de tendências até à construção de modelos multivariados.

Inicialmente a previsão das audiências consistia em assumir que o futuro seria igual ao passado recente e que o erro em que se incorre por utilizar a média dos resultados do último ano seria desprezível. Esta abordagem a nível nacional ainda se utiliza, apesar do crescente número de ofertas televisivas. A nível internacional, segundo a bibliografia da especialidade, em geral já se encontra fortemente ultrapassada.

Com base nos valores históricos observados podemos recorrer à projecção de tendências e ao cálculo dos coeficientes de estacionalidade para prever os valores de *ratings* ou de outros indicadores. Estas abordagens podem apresentar crescentes limitações decorrentes das mudanças verificadas no sector audiovisual, resultante da sua desregulamentação, assim como da sucessiva fragmentação das audiências.

Estes sistemas de previsão devem produzir melhores resultados em contextos onde o painel tenha acesso a um reduzido número de canais e onde a audiência semana após semana não varie de forma significativa o seu perfil de consumo televisivo. A actual estrutura do sector, em países mais desenvolvidos, corresponde à multiplicação das opções televisivas, e a uma maior irregularidade do consumo o que poderá invalidar a aplicação destas abordagens. Porém, a nível nacional persiste a dúvida sobre a sua actual validade, uma vez que uma significativa percentagem do painel apenas tem acesso aos principais canais generalistas. Talvez seja viável recorrer a estas análises para prever os valores das taxas médias de audiências para o total da televisão. Algumas análises a este nível constam do *Apêndice F.1*, mas uma vez que se considera um período característico a nível médio do consumo televisivo, não se identificam tendências, mas, sim, comparam-se diferentes abordagens de previsão não linear que possam viabilizar a sua previsão.

Contudo, alguns autores internacionais acabam por considera-las como pouco relevantes uma vez que a explicação da sua evolução não se encontra na série em si mesma, mas, sim, nas variáveis exógenas que não se relacionam com a tendência observada (*Danaher et al., 2011*). Estas abordagens devem ser consideradas quando se pretende a previsão dos indicadores que se encontram menos dependentes da evolução verificada ao nível da grelha de programas e no sector, como, por exemplo, para as taxas médias de audiências para o total da TV. Segundo *Bunge (1981)*, podemos apresentar alguma confiança nas tendencias analisadas teoricamente desde que conhecendo os mecanismos causantes das suas principais evoluções e tendo algumas razões para supor que o futuro será parecido ao passado.

Mas, quando não existe nenhuma mensagem e não nos encontramos com uma linha realmente de tendência apenas será possível realizar uma vaga previsão estatística, referente à sequência como um todo e não em relação às suas componentes. “Las líneas de tendencia son, como de proyección, dignas de escasa confianza si no se basan en alguna teoría” (*Bunge, 1981, p.638*).

De acordo com *Bunge* a previsibilidade depende do nosso conhecimento sobre as leis e, logo, as análises explicativas são essenciais. Frequentemente, a imprevisibilidade decorre da ausência de uma teoria em sentido próprio em determinada área de estudo. Salienta que os acontecimentos causais são individualmente imprevisíveis, mas que a imprevisibilidade não significa sempre um carácter causal e que a causalidade individual leva a regularidades colectivas. A causalidade e a projectabilidade são relativas e dependem do corpo de conhecimento disponível em determinada área.

Os métodos univariados estritamente extrapolativos limitam a sua utilização directa a ambientes mais estáveis. Apesar do consumo televisivo poder ser mais variável face ao que se verificava quando apenas existia a televisão pública, a adequabilidade destas abordagens depende da variabilidade dos indicadores agregados e que poderá não ser representativo do que se verifica a nível individual. Estas metodologias podem ser utilizados para a previsão de factores a utilizar nos métodos causais e no estabelecimento de tendências em análises de sensibilidade.

Na previsão das audiências, também, podemos utilizar análises matemáticas que inter-relacionem e ponderem factores previamente identificados, assim como recorrer a sistemas de simulação que permitam identificar regularidades estatísticas na evolução dos dados. Nos modelos causais destacam-se as análises de Regressão Linear, não tendo sido identificados na bibliografia modelos de Regressão Dinâmica e de transferência que permitem analisar a relação dinâmica entre as séries temporais, assim como os Métodos Bayesianos.

Em análises multivariadas, os factores a considerar devem ser mais fáceis de prever do que a variável dependente ou serem conhecidos. Porém, é muito difícil identificar variáveis explicativas da evolução dos dados que, também, apresentem capacidade predictiva considerando, quer as variáveis de audimetria, quer dados mais desagregados de caracterização da exposição televisiva.

A extracção de regularidade estatísticas, por parte da oferta e da procura é difícil de realizar. Difícilmente será viável identificar Regras que permaneçam, a médio prazo, válidas quando incidem nos comportamentos sociais, e consequentemente, também, nos comportamentos de consumo televisivo. A inadequabilidade das metodologias existentes até ao desenvolvimento de metodologias não proposicionais de *Data Mining* também contribui para uma maior dificuldade na sua identificação através da análise das bases de dados de audimetria que apresentam uma elevada dimensionalidade. O desenvolvimento dos sistemas expertos também contribuiu para uma maior facilidade na identificação de regularidades na evolução dos indicadores de audimetria. Os avanços observados ao nível da informática, especialmente na Inteligência Artificial e em Lógica Difusa e as inúmeras metodologias de DM que existem actualmente tornam esta perspectiva de análise numa das mais promissoras.

Os Sistemas Expertos procuram imitar o funcionamento da inteligência humana, no sentido de poder valorizar de forma diferente novas informações, inclusive eliminando as que não são relevantes e modificando o seu comportamento em função da experiência. Esta tecnologia também tenta produzir previsões mediante algoritmos matemáticos, mas modificando-as em função de juízos de valor e experiências subjectivas e tendo em conta o dia, canal e hora de emissão. Na bibliografia da especialidade, esta opção é apresentada como a mais interessante e que poderá produzir os melhores resultados no futuro. Consideram sistemas de aprendizagem de máquinas, como, por exemplo, as Redes Neurais onde se simula uma espécie de cérebro que consiga resolver todas as situações que possam surgir no âmbito da exposição televisiva com um elevado grau de acertos. Para a Rede Neuronal tomar decisões será necessário treina-la apresentando-lhe exemplos históricos com as respectivas soluções.

III.3. CONSIDERAÇÕES SOBRE AS ANÁLISES DE PREVISÃO NAS CIÊNCIAS SOCIAIS: AS CARACTERÍSTICAS DOS MODELOS DE EXPOSIÇÃO

O planeamento de *media* poderá, eventualmente, ser realizado considerando diferentes perspectivas e níveis de agregação dos dados. Porém, a previsão dos comportamentos individuais em qualquer área levanta inúmeras dificuldades. A evolução dos índices de audiência resulta dos comportamentos de milhares de telespectadores que apresentam a nível individual aleatoriedade e imprevisibilidade. Segundo *Moritz e Pereira (2006)*, entre muitos outros autores, os comportamentos individuais são indeterminados, acausais e imprevisíveis:

A acção humana é indeterminada, variada e depende das condições do momento de escolha ou da ocasião em que a decisão é tomada. Factores emocionais influenciam a tomada de decisão. Muito desses elementos são inconsistentes, mas estruturam a percepção humana em dado momento, influenciando a decisão. Sob condições diferentes, em outro ambiente ou sob outros tipos de pressões emocionais, as escolhas dos mesmos indivíduos podem transformar-se e suas acções serem diferentes. (p.65)

Alguma bibliografia, entre os quais, *Bukharin (1933)*, refere que não se pode prever a data em que se realizará um acontecimento, não conhecemos as leis da evolução social ao ponto de as podermos exprimir em algarismos exactos e ignoramos a velocidade dos processos sociais, mas podemos indicar a sua direcção mais provável. Verificámos que, pontualmente, alguma bibliografia defende a previsibilidade dos dados nas Ciências Sociais, mas apenas ao nível de tendências nos comportamentos agregados e, não, relativamente a condutas específicas no espaço e tempo. Consequentemente, por vezes, nas análises de previsão realizadas a um nível mais desagregado, apenas se avalia os factores que podem eventualmente condicionar determinado resultado, independentemente de quando este possa ocorrer e atribuindo-lhe uma determinada probabilidade .

Recorrer ao histórico para prever tendências nos comportamentos individuais, também, poderá não conduzir a resultados satisfatórios. De acordo com *Bulck (2006, p.250)*, a exposição televisiva individual verificada em determinado instante não apresenta uma relação e dependência face à exposição num momento futuro, inviabilizando a sua previsão: “Exposure at Time 1 only explains an additional 7% of news exposure at Time 2 ... This data also suggests that for explaining or predicting news viewing follow up studies may not be very productive.”

Nas Ciências Sociais imperam as análises explicativas, uma vez que é mais fácil analisar e compreender um mecanismo do que conseguir estabelecer uma relação de causa e efeito, devido à repercussão de variáveis, frequentemente, não quantificáveis e previsíveis. O comportamento individual é classificado como imprevisível, indeterminista, estocástico no qual apenas probabilidades podem ser determinadas (*Robertson, Zielinski & Ward, 1984*). Apenas se poderá avaliar a propensão para se verificarem determinados comportamentos no decorrer do tempo. Inevitavelmente, as análises agregadas para o universo televisivo ou considerando segmentos devem ser privilegiadas quando se pretende realizar a previsão da exposição aos suportes publicitários com as menores margens de erro.

A exposição pode ser quantificada recorrendo aos índices de audiências ou a valores de fluxos de audiências, ou seja, a uma percentagem ou número de telespectadores. As eventuais regularidades manifestam-se, essencialmente, num colectivo de indivíduos, e, não, a nível individual. Se existirem regularidades nos comportamentos, apenas será possível a sua identificação recorrendo a um grande horizonte temporal e a grandes bases de dados que possam determinar a estrutura geral do comportamento dos telespectadores. Os trabalhos desenvolvidos na área da duplicação das audiências entre os suportes publicitários pretende avaliar a regularidade dos comportamentos, entre outros objectivos.

Os conteúdos seleccionados pelos telespectadores não correspondem, necessariamente, ao que estes últimos manifestam como sendo da sua preferência. Devemos identificar os factores e os contextos que determinam a exposição televisiva e influenciam em maior grau a probabilidade de exposição a determinado evento. Verificamos que existem múltiplas perspectivas explicativas que na realidade não se encontram integradas e que apenas proporcionam uma visão parcial, incompleta e subjectiva dos comportamentos de consumo televisivo. As sondagens de intenção de ver, entre outras abordagens qualitativas de previsão das audiências que também se fundamentam nas preferências e na eventual atractividade que os suportes exercem no público têm apresentado no passado grandes margens de erro.

Segundo *Hauser (1978)*, os modelos de preferência também não têm como função prever os comportamentos de forma “determinista”, mas, sim, explicar e prever o máximo que for possível, dos comportamentos observados. Devemos adoptar um raciocínio idêntico quando se analisam os valores agregados. O autor refere, que a nível individual, é necessário combinar o poder dedutivo das teorias de utilidade com os modelos estocásticos dos comportamentos da econometria e com os poderes de medição dos comportamentos do consumidor.

Os modelos tradicionais do Marketing, também, tendem a considerar os consumidores numa perspectiva agregada e média (*Patel & Schlijper, 2008*). Idealmente seria desejável relacionar o comportamento do ponto de vista do consumidor com a perspectiva do gestor de uma marca, assim como a exposição agregada e a evolução dos índices de audiências que condicionam a venda do espaço comercial.

Em suma, as análises de previsão não devem ser realizadas com base num interesse, preferência ou comportamento passado, mas, sim, recorrendo à interpretação da relação que existe entre um colectivo de individuos e um conjunto de suportes ou aos indicadores de audimetria e em especial aos valores de *ratings*.

Segundo *Serra (2005)*, a perspectiva da imprevisibilidade e acausalidade dos comportamentos humanos individuais e a dificuldade de realizar previsões a nível agregado contrasta com a óptica que impera nas Ciências Exactas:

As palavras fascinantes da nova ciência, tais como caos e indeterminismo, transformadas em símbolos por alguns autores, não servem para caracterizar o saber nas Ciências Exactas ... A ideologia dominante dos que trabalham em Ciências Exactas é sem dúvida o determinismo ... Quando investiga ele confia na existência de uma “ordem”, ou seja, ele procura regularidades, periodicidades, estruturas, princípios de conservação e leis que se exprimem matematicamente. Se tal não fosse possível, não haveria trabalho para um cientista. (*p.6, p.7*)

O **determinismo**, a causalidade e a previsibilidade também não são conceitos equivalentes (*Filho, 2008*). O caos determinista corresponde à forte influência das condições iniciais sobre estados futuros em sistemas deterministas, inviabilizando as previsões. O caos não constitui um argumento contra o determinismo, mas mostra que a previsibilidade como sonhada por Laplace nem sempre é concretizável (*Silveira, 1993*). Por outro lado, a previsibilidade poderá ser viável em contextos classificados como “caóticos”: “... mesmo trabalhando com os fenómenos “indeterministas” da Mecânica Quântica, ou os “caóticos” da Dinâmica, é possível fazer determinadas previsões e construir aplicações.” (*Serra, 2005, p.6*).

A causalidade, também, não se reduz à capacidade de prever e nem causalidade se reduz a determinismo e deve-se recorrer à interpretação de probabilidade enquanto propensão para se verificar determinado comportamento (*Filho, 2008, p.11*): “Podemos admitir um mundo indeterminista, aberto a vários cenários de possibilidades que constituem propensões reais”. Podemos conceber um mundo quântico causal e indeterminista no qual vigorem as propensões de Popper (*Filho & Selleri, 2006*).

De acordo com *Taleb e Makridakis (2009)*, existe uma grande resistência em aceitar de que em inúmeros domínios da Economia, Finanças e Gestão as previsões não são realmente viáveis apesar das evidências empíricas e, que consequentemente, devemos adoptar outras perspectivas de análise:

The majority of academics and business people were not willing to accept these findings, and instead preferred the illusion of control, pretending that accurate forecasting was possible. ... a large number of decision makers and individuals believe that accurate forecasting is possible, and that uncertainty can be assessed correctly and controlled effectively. ... we must realize that although accurate forecasting is possible in some areas of physics and engineering, it fails when complex systems are involved. Moreover, such a failure becomes more pronounced in complex social systems where the actions and reactions of people can and do affect future outcomes. (p.841, p.842)

O homem, apesar das Regras sociais e de outro tipo de restrições a que se encontra sujeito, goza de livre arbítrio, ou seja, tem o poder de escolher as suas acções. O livre arbítrio implica que uma acção não é completamente condicionada por factores antecedentes, e a percepção de que a acção originou-se na sua vontade. O livre arbítrio corresponde a uma propriedade biológica, um traço de personalidade segundo o qual o cérebro possui a liberdade de gerar comportamentos e opções.

Inclusive no mundo animal foi estabelecido há muito tempo que o “comportamento determinístico”, a ideia que um animal provocado de tal forma vai reagir sempre com a mesma resposta não corresponde a uma descrição completa do seu comportamento. Os animais não são totalmente previsíveis, mas as respostas podem ser enquadradas em um conjunto fixo de opções, às quais podemos associar um conjunto definido de probabilidades. A variabilidade que é inerente ao seu comportamento é um pré-requisito para a sobrevivência em um ambiente competitivo.

Uma das coisas mais deprimentes da adopção determinista é a que comporta descrever um mundo insuportavelmente indutivo, no qual o passado é igual ao futuro e no qual não há nada de novo e criativo. ... Se tudo já fosse determinado de tal maneira que qualquer coisa que viéssemos a fazer e a pensar fosse de antemão previsível, então a nossa própria dignidade e com esta os valores de liberdade, de autonomia e de livre arbítrio seriam uma ilusão ou estariam seriamente comprometidos. (Filho, 2008, p.3304-2)

O consumo televisivo encontra-se, sem dúvida, condicionado por inúmeras variáveis algumas das quais não quantificáveis ou imprevisíveis, como, por exemplo, a climatologia, determinadas características dos conteúdos transmitidos, a repercussão da imagem dos canais, as estratégias de programação e de contra-programação, os factores temporais, a disponibilidade para ver TV e de predisposição de exposição ao *media*, a

influência do agregado familiar, o calendário escolar e laboral, preferências, expectativas, interesses, motivações, vontade, atitudes, acesso ao *media*, o conhecimento da programação, a predisposição para se desenvolverem comportamentos estruturais de consumo televisivo e por características psicográficas, económicas e culturais da população.

Por exemplo, em *Vandebosch, Roe e Van den Bulck (2006)* avalia-se a influência dos ciclos lunares no consumo televisivo e os resultados são inconsistentes, existindo alguma repercussão das variáveis consideradas nas análises:

The range of postulated lunar effects encompasses both physiological and psychological / behavioural concerns ... Future studies should investigate data over a longer period of time, include additional moon variables (i.e., amount of moonlight), explicitly test the moon-sleep derivation-television viewing hypothesis, control for other environmental factors (e.g., weather conditions), and search for alternative explanations for a possible relation between amount of television viewing and the moon cycle. (p.287)

Em *Mares, Oliver e Cantor (2008)*, salienta-se a repercussão dos estados emocionais diários dos telespectadores nos suportes seleccionados por parte de diversas faixas etárias e a necessidade de avaliar a repercussão da exposição nesses mesmos estados emocionais. A selecção de determinado suporte condiciona as futuras selecções, assim como os estados emocionais presentes e futuros o que contribui para a imprevisibilidade dos comportamentos de consumo televisivo.

Frequentemente, encontramos na bibliografia da especialidade, assim como em muitas outras áreas das ciências sociais, a designação de análises de previsão quando na realidade apenas pretendem avaliar os factores que podem, por vezes, condicionar a exposição em determinado contexto, mas sem conseguir prever acções concretas no espaço e no tempo. Isto decorre de nas Ciências Sociais se assumir como inadequado as análises de previsão com um carácter “determinista”, sendo apenas possível indicar tendências ao nível da evolução dos dados.

Porém, eventualmente, da agregação de comportamentos individuais definidos maioritariamente nas Ciências Sociais como imprevisíveis e acausais podem resultar comportamentos não caóticos e verificar-se alguma regularidade a nível colectivo, viabilizando assim a previsão de alguns índices de audiências e que será função dos contextos em análise. Uma vez que o todo não corresponde ao somatório das partes, por vezes as tendências individuais complementam-se podendo incutir uma maior estabilidade aos dados e que resulta imprescindível para a previsão dos índices de audiências, observando-se um certo “determinismo estatístico”.

O fenómeno social não depende da natureza dos indivíduos. É que na fusão que deles resulta, todos os caracteres individuais, divergentes por definição, se neutralizam e se dissipam mutuamente. ... O resultado não se produziria se elas o excluíssem, mas não são elas que determinam esse mesmo resultado. (*Durkheim, 1982, p.198*)

Um fenómeno social apresenta propriedades colectivas e a indução implica a necessidade de se observarem regularidades nos acontecimentos observados que eventualmente se podem manifestar no colectivo. O grau de regularidade na exposição televisiva condiciona a construção dos planos de *media* e pode ser avaliada através dos estudos de duplicação de audiências (*Henriksen, 1985*) nas suas diversas vertentes.

As análises de previsão mais adequadas não implicam necessariamente a utilização de sofisticados métodos de análise e devem ser privilegiadas as análises estatísticas em detrimento das previsões mais qualitativas recorrendo a juízos de valor, que por vezes são utilizadas por parte das estações televisivas, mas em menor grau e apenas como complemento numa perspectiva de construção de planos de *media*. O facto de dificilmente serem identificados contextos históricos equivalentes a um momento presente também invalida a aplicação de determinadas abordagens que promovem a identificação de cenários comparáveis à observação em análise.

Statistically sophisticated or complex models fit past data well but do not necessarily predict the future accurately – meaning that they have been data-mined. ... Simple models do not necessarily fit past data well but predict the future better than complex ones or sophisticated statistical models. ... Ironically in most domains, judgmental predictions are less accurate than those of statistical models, as they are influenced by human biases and limitations. ...The future isn't what it used to be, or alternatively, history never repeats itself in exactly the same way. (*Taleb & Makridakis, 2009, p.841*)

Quando falamos em previsão tal conceito é demasiado abrangente, uma vez que existe uma grande diversidade de contextos e cada um deles exige diferenciadas abordagens analíticas e fontes de dados, em função da variabilidade observada nos dados e da dimensão dos erros de previsão.

Os contextos dependem, por exemplo, do público objectivo, faixas horárias, indicadores de audimetria em análise e nível de agregação dos dados, o canal, o mês, as características da grelha de programação, as estratégias de programação e de contra-programação, o planeamento comercial, as tipologias programáticas, as variáveis sociodemográficas, entre muitas outras. No Reino Unido, foi realizada uma tese de Doutoramento apenas para tentar “prever as audiências” dos jogos de futebol (*Buraimo, 2007*) e que na realidade avalia em maior grau os factores que condicionam a sua evolução e uma previsão direccional.

Atendendo ao exposto anteriormente, a previsão da exposição televisiva considerando o comportamento individual, dificilmente, será viável e devemos avaliar a regularidade da exposição a nível colectivo, assim como as potencialidades de novas áreas metodológicas, uma vez que imperam na bibliografia da especialidade as mais tradicionais (*Weber, 2002*). Não será viável realizar previsões exactas, mas por menores que sejam os ganhos ao nível da precisão, estes podem representar grandes poupanças no investimento realizado em publicidade televisiva.

Anteriormente, quando existia um reduzido número de canais e poucas opções ao nível dos *media*, o desenvolvimento de modelos de previsão e explicativos das audiências era mais facilmente exequível e a reduzida bibliografia da especialidade incide nestes contextos: "All past *ratings* forecasting research has examined markets with between 3 and 6 channels, with only one study using data from the 21st century." (*Danaher et al., 2011, p.2*)

Um plano de *media* é construído com base na previsão de inúmeros indicadores e que se baseia na cobertura, repetição e qualidade da comunicação. Porém, a compreensão do alvo ao nível do seu perfil de consumo e exposição aos *media* também é determinante, nomeadamente por avaliação das influências estruturais. Trata-se de um conjunto de metodologias apresentam como finalidade o conhecimento das características do receptor que forma parte do processo de comunicação, sendo esta uma informação indispensável para que os planificadores possam tomar decisões fiáveis, contribuindo, assim, para uma maior eficácia das estratégias de comunicação publicitárias.

Por vezes, os objectivos explicativos e de previsão confundem-se na bibliografia da especialidade, atendendo às dificuldades que a previsão coloca neste domínio, mas na realidade não são equivalentes, ainda que apresentem uma forte interdependência. Os modelos explicativos diferenciam-se das análises de previsão a vários níveis. Por exemplo, podem incluir variáveis que condicionam a evolução dos indicadores, mas não sendo previsíveis, quantificáveis ou conhecidas no momento de realizar a previsão (Weber, 2002). Uma maior complexidade do modelo pode levar a um maior ajustamento à base de dados de treino e alcançar uma maior capacidade explicativa da evolução dos dados, mas também poderá contribuir para uma menor capacidade predictiva quando se pretende a quantificação de uma variável num contexto concreto no espaço e tempo (Hastie, Tibshirani & Friedman, 2001).

As análises de previsão, frequentemente, apenas tentam “prever” a probabilidade de ocorrência de um sucesso no decorrer do tempo. As variáveis a considerar podem introduzir ruído e reduzir a precisão das análises se tiverem que ser anteriormente previstas. Porém, as análises explicativas também contribuem e condicionam fortemente a selecção das abordagens metodológicas que forem mais adequadas à previsão, assim como a avaliação da sua viabilidade. Apesar da relevância de ambas as perspectivas, na bibliografia das Ciências Sociais existe uma maior ênfase na vertente explicativa, ainda que também escassa quando se pretende a análise de “*ratings*”.

III.4. O CONTRIBUTO DOS MODELOS EXPLICATIVOS DO CONSUMO TELEVISIVO E DOS RATINGS PARA A CONSTRUÇÃO DOS MODELOS DE EXPOSIÇÃO

A maturidade de uma área apresenta os fundamentos da sua previsão, ao nível da sua viabilidade e da identificação das perspectivas que podem conduzir aos melhores resultados. O comportamento do público é considerado como sendo cada vez mais complexo e dinâmico (*Weber, 2002*):

The increasing need for quantitative models suitable for explaining and forecasting TV usage in order to accommodate the growing complexity of audience behaviour has also been noted from an academic point of view. ... The complexity of this behaviour however is growing more rapidly than our analysing methods. ... Only the development of new quantitative research methods will allow us to understand television usage in the future ... The available literature suggests that only few attempts have been made to model or forecast television-viewing patterns for target audiences by employing sophisticated methods of analysis. ... Searching the Internet (there are almost no studies available in public libraries) yields only a few commercial projects inside and outside Europe looking onto the explicability and predictability of *ratings* by means of Artificial Intelligence. ... Communication sciences are focused on explanatory (as opposed to forecasting) models. Models directed towards explaining TV usage are also well suited for television audience prediction as they supply information on possible determinants of TV use in general. (*p.1*).

Permanecem inúmeras incógnitas relacionadas com os factores que condicionam maioritariamente a exposição televisiva e a evolução dos “*ratings*”, assim como as respectivas interacções (*Cooper & Tang, 2009; Patelis et al., 2003; Webster, 2006*). O facto de imperar na bibliografia da especialidade a aplicação de abordagens de análise estatística de dados e fontes de dados mais tradicionais também tem penalizado a investigação realizada. Verificamos que existe a necessidade de avaliar o contributo de áreas mais recentes para a segmentação do painel de audimetria, para a selecção dos suportes a incluir em planos de *media* e para a realização de análises explicativas e/ou de previsão dos “*ratings*”.

Actualmente, poderá existir uma maior irregularidade e variabilidade na exposição, o que requer a aplicação de métodos de análise estatística mais adequados e das fontes de dados que viabilizem uma análise mais precisa, evolutiva e detalhada da exposição.

A suposição de existência de regularidade do consumo televisivo resulta dos métodos de investigação mais tradicionais que recorrem aos inquéritos, mas que na realidade são incapazes de prever as audiências televisivas e partem de pressupostos que são inválidos.

A caracterização da exposição televisiva e a compreensão dos factores que a influenciam será essencial para conseguir melhores resultados na área da quantificação das audiências televisivas futuras. A forma como os telespectadores interagem maioritariamente com o meio televisivo e os programas condiciona a selecção das abordagens de previsão mais adequadas para diferentes contextos. Inevitavelmente, previamente será necessário desenvolver análises exploratórias e explicativas da evolução dos dados.

Segundo *Patelis et al. (2003)*, as análises exploratórias e explicativas são necessárias para que se possa definir a tipologia de modelos quantitativos a utilizar, posteriormente, em análises de previsão. O consumo televisivo apresenta uma grande complexidade e os factores que o determinam ainda não estão bem definidos e hierarquizados. A previsão pode implicar e depender da caracterização e do desenvolvimento de modelos explicativos que permitam interpretar a estrutura dos comportamentos. Os melhores resultados do ponto de vista da previsão só podem ser alcançados após uma exaustiva caracterização e explicação dos comportamentos de consumo televisivo e da evolução dos indicadores de audimetria, num sector que é muito dinâmico e fortemente influenciado pelo horizonte temporal e geográfico em análise.

III.4.1. Os modelos explicativos da exposição televisiva

Os modelos de “*ratings*” têm sido associados a uma das perspectivas explicativas da exposição televisiva e que segundo a bibliografia da especialidade apresenta uma maior capacidade predictiva: a perspectiva estrutural. A viabilidade da previsão da exposição também depende dos factores que a condicionam maioritariamente.

Os planos de *media* podem ser construídos através da previsão dos “*ratings*” ou dos fluxos de audiências entre os suportes publicitários. Consequentemente, no *Apêndice D* consta uma caracterização mais exaustiva dos modelos explicativos da exposição televisiva. Neste capítulo pretende-se apenas enquadrar a investigação que recorre às bases de dados de audimetria.

Nesta área predominam teorias e modelos de caracterização e explicação da exposição televisiva, frequentemente, contraditórios entre si. Os trabalhos existentes podem ser agregados, essencialmente, em duas vertentes de análise que, desde há longa data, se confrontam, uma vez que perseguem objectivos diferentes: a abordagem das gratificações e usos do *media* e a perspectiva da teoria estrutural. A diversidade da formação de base dos investigadores, por exemplo, nas áreas da Psicologia, Sociologia, Economia, Estatística, Gestão, Antropologia, Comunicação Social, etc., e ao nível dos contextos sócio-culturais e económicos dos países em que os estudos são desenvolvidos contribuem para esta realidade.

A teoria de gratificações e usos do *media* (Rubin, 1993) pode ser associada a uma vertente de análise mais qualitativa no âmbito das Ciências Sociais. Corresponde à perspectiva que impera a nível Europeu e na América Latina (Huertas, 2002). Os autores com base em inquéritos e em entrevistas ou recorrendo a métodos mais qualitativos defendem o carácter activo do público, sendo a exposição função das preferências e da satisfação de necessidades subjacentes.

A actividade do público é descrita como um sinónimo de utilidade, selectividade e intencionalidade, desempenhando um papel importante no processo de comunicação. Frequentemente, ignoram que existem constrangimentos no consumo televisivo.

Porém, identificamos alguns autores que defendem que a exposição ao *media*, independentemente das escolhas realizadas, pode apresentar uma maior passividade e decorrer de hábitos desenvolvidos no decorrer do tempo ou de outros factores estruturais (Bryant & Zillmann, 1985; Ronsenbegren, Wenner & Palmgreen, 1985). As sondagens de intenção de ver, nas quais os telespectadores manifestam os seus interesses e preferências, apresentam grandes margens de erro quando se utilizam como método de previsão das audiências.

A teoria dos factores estruturais (Cooper, 1993; Cooper, 1996) explica melhor o comportamento do público quando imperam as motivações ritualísticas, ou seja, quando existe uma maior relação para com o *media*.

A perspectiva estrutural defende que os telespectadores seleccionam os conteúdos passivamente, segundo hábitos desenvolvidos no decorrer do tempo e aderem aos constrangimentos estruturais, como, por exemplo, os que decorrem dos factores de planeamento. Sob esta perspectiva, a exposição televisiva será função de factores de disponibilidade do painel, acesso ao *media*, rotinas desenvolvidas, estruturação do mercado, entre outros factores. Os telespectadores são designados, essencialmente, como “orientados para o *media*”, uma vez que recorrem à televisão para preencher os seus tempos livres ou “vazios”, ou seja, um não tempo, seleccionando os melhores conteúdos disponíveis nesse momento. Porém, a selecção também poderá não corresponder aos conteúdos que eventualmente suscitam um maior interesse em função das alternativas disponíveis em determinado momento. Neste contexto, predominam as motivações ritualísticas que reflectem utilidade, mas uma menor intenção, selectividade e atenção.

As duas perspectivas funcionam a diferentes níveis e perseguem diferentes objectivos. Na actualidade, não existe consenso quanto à perspectiva que maioritariamente influencia a exposição televisiva e quanto à sua actual validade e respectiva interacção (Cooper & Tang, 2009). De acordo com alguns autores, as duas ópticas influenciam os padrões de consumo televisivo (Cooper, 1993), como, por exemplo, em *Comstocks* (1980) onde se refere que a decisão de ver televisão poderá apresentar uma maior passividade, mas a escolha do que ver poderá ser mais activa. Este processo a duas fases sublinha a debilidade de considerar as preferências como o único componente da previsão do consumo televisivo.

Atendendo às divergência dos pressupostos em que se baseia a teoria de gratificações e usos do *media* e a teoria estrutural verificamos que existe um reduzido número de publicações que promovem a sua integração e a avaliação das respectivas interdependências. Podemos citar o modelo teórico proposto em Webster e Wakshalag (1983) e o trabalho empirico descrito em Cooper e Tang (2009) que incide na análise da exposição ao *media* ou em Wonneberger, Schoenbach e Van Meurs (2012) onde se analisa o processo de selecção dos espaços informativos.

Ambos os trabalhos empiricos referidos concluem que nos contextos em análise prevalece a influência estrutural e que são escassas as interacções observadas entre as variáveis estruturais e as individuais, recorrendo o primeiro trabalho a um inquérito e o segundo às bases de dados de audimetria.

Cooper e Tang (2009) classificam os telespectadores como “activos em ambientes estruturais”. O comportamento da audiência apresenta uma crescente complexidade, quer a nível individual, quer quando se consideram segmentos que são dificilmente identificáveis (Kim, 2001). De acordo com Copper e Tang (2009) as características sociodemográficas condicionam em maior grau a disponibilidade da audiência e não a selecção. Defendem a necessidade de recorrer a variáveis mais discriminatórias dos comportamentos de consumo televisivo.

Contudo, as variáveis sociais e demográficas *a priori* podem condicionar a evolução dos indicadores de audimetria, uma vez que o grau de instrução, a capacidade económica, a região, a faixa etária e o sexo podem influenciar a frequência, horários, preferências e quais os programas e as tipologias mais seleccionados. Também, poderá condicionar a influência dos factores estruturais de planeamento nos comportamentos de consumo televisivo.

Na opinião de *Cooper e Tang (2009)*, a exposição televisiva resulta em maior grau do hábito, acesso, motivações ritualísticas e dos factores de disponibilidade, entre outras variáveis estruturais. Porém, o hábito também não implica necessariamente uma maior regularidade.

Os autores tentaram avaliar a repercussão de variáveis individuais e estruturais, e concluíram que a vertente estrutural permanecerá no futuro como uma das principais condicionantes da exposição televisiva. Por outro lado, a perspectiva estrutural não estará necessariamente associada a comportamentos mais passivos dos telespectadores. Também, referem o confronto que existe entre a investigação quantitativa e a qualitativa das audiências televisivas e as inúmeras incógnitas que permanecem, hoje em dia, relativamente ao consumo televisivo, assim como a necessidade de novas abordagens e fontes de dados quantitativas serem adoptadas. Os inquéritos ou inclusive os diários de audiências apresentam inúmeras limitações para a caracterização e previsão da exposição.

De acordo com *Wonneberger et al. (2012)*, por análise de duas décadas de dados de audimetria holandeses que caracterizam a exposição aos espaços informativos, a crescente exposição a esta tipologia de programas encontra-se maioritariamente influenciada pelos factores estruturais e de contexto, e em muito menor grau por parte das motivações dos telespectadores. Os mesmos autores, em *Wonneberger, Schoenbach e Van Meurs (2011)*, também por análise da exposição aos espaços informativos concluem, uma vez mais, que os factores de contexto prevalecem largamente sobre os factores motivacionais, representando um efeito sete vezes superior.

Em suma, a perspectiva estrutural e a de gratificações e usos do *media* funcionam a diferentes níveis de análise e apresentam objectivos radicalmente diferentes. Actualmente, persistem inúmeras incógnitas quanto aos factores que condicionam maioritariamente a exposição televisiva e os indicadores que dela resultam e que podem corresponder a variáveis individuais ou estruturais, sendo especialmente difícil a sua integração. Consequentemente, são raros os contributos existentes, a nível teórico e empírico, para o desenvolvimento de modelos integrativos de análise da exposição televisiva.

III.4.2. A análise dos comportamentos estruturais

No sector televisivo existe a necessidade de definir o conjunto de suportes que possa garantir uma frequência de contacto que seja suficiente para incentivar a aquisição dos produtos e serviços divulgados. Desta necessidade surgiu o desenvolvimento das análises de duplicação das audiências, segundo a perspectiva estrutural. De acordo com a bibliografia da especialidade, os comportamentos estruturais apresentam uma maior capacidade predictiva e utilidade para explicar o comportamento agregado (*Cooper, 1996*).

A investigação realizada ao nível da análise das audiências conjuntas abrange quatro vertentes que avaliam a tendência que se observa nos telespectadores para estar presente na audiência de um conjunto de suportes.

Consequentemente, relaciona-se com a previsão das audiências, através da avaliação da regularidade da exposição e dos factores que a condicionam, assim como com a avaliação da eficácia das estratégias de programação e de contra-programação que pretendem a fidelização das audiências e a maximização da exposição. Também contribuem para a estruturação dos planos de *media* onde se pretende seleccionar um conjunto de suportes que permitam alcançar uma maior cobertura do mercado alvo, mas também alguma repetição de exposições para que os produtos e serviços possam ser adquiridos.

Uma elevada “brand equity” de determinado suporte está directamente relacionada com um comportamento de lealdade por parte das suas audiências, o que aumenta, por conseguinte, a estabilidade e previsibilidade do alcance desse suporte. Os modelos de previsão devem prever, não só, o nível de audiências de um programa, mas também os hábitos das audiências onde se inserem as inserções publicitárias e que corresponde à duplicação das audiências. Os modelos de exposição, em alternativa aos *ratings*, também têm como função quantificar a distribuição de frequências dos contactos alcançados por um plano.

Segundo *Henriksen (1985)*, o conceito de duplicação das audiências engloba três componentes designadas de lealdade ao canal, efeitos de herança entre programas adjacentes e a lealdade programática. A lealdade aos canais avalia a tendência dos telespectadores para seleccionar programas de determinado canal (*Cooper, 1993*). Os efeitos de herança representam a análise dos níveis de duplicação de audiências entre programas adjacentes ou próximos num mesmo canal e dia (*Boemer, 1987; Cooper, 1993; Davis & Walker, 1990; Walker, 1988; Webster, 1985*). A lealdade a programas incide na análise dos níveis de audiência conjunta entre diferentes episódios de determinado programa (*Barwise, 1986; Barwise, Ehrenberg & Goodhardt, 1982; Headen, Klompmaker & Rust, 1979; Sherman, 1995; Webster & Wang, 1992*). Também se poderá incluir nesta área a análise da exposição repetida que representa a exposição a retransmissões de um programa. Actualmente, não existem dúvidas de que as características de planeamento da grelha de programas produzem determinadas tipologias de comportamentos, tornando complexa a relação existente entre as preferências e a selecção que realmente se verifica (*Webster & Litchy, 1991*).

Os trabalhos de duplicação de audiências são desenvolvidos sob a perspectiva explicativa estrutural e recorrem, essencialmente, a fontes de dados de audimetria incidindo uma vez mais em trabalhos realizados em países anglo-saxónicos. As conclusões não são generalizáveis de país para país, atendendo às diferenças culturais, económicas, sociais e de estruturação do sector.

Verificamos uma escassez de trabalhos publicados neste âmbito a nível Europeu e em outros países não anglo-saxónicos, e, logo, por vezes, surgem análises comparativas face aos resultados apresentados, essencialmente, por Britânicos e Norte Americanos, como, por exemplo, em *Zubayr (1999)*.

Corresponde a mais uma das áreas onde a nível nacional não foi viável identificar trabalhos publicados, assim como na área dos modelos explicativos de “*ratings*”. A exposição televisiva poderá não corresponder a preferências da audiência, assim como decorrer de um processo intencional que visa satisfazer necessidades e interesses subjacentes, sendo esta a perspectiva da abordagem das gratificações e usos do *media*.

Em suma, as análises de duplicação contribuem para o planeamento de *media* a diversos níveis e traduzem os factores e comportamentos associados à estrutura das opções programáticas disponíveis. As análises de duplicação referem-se aos “comportamentos estruturais” uma vez que resultam dos factores de planeamento e contribuem para a avaliação da regularidade da exposição.

III.4.3. O conceito de hábito de consumo televisivo

O conceito de “hábito” corresponde a uma temática que suscita alguma controvérsia, uma vez que em geral é identificado como uma indicação da passividade da audiência. Porém, também tem sido operacionalizado como uma das possíveis gratificações do acesso ao *media*. Também poderá não ser uma mera tipologia de “usos do *media*”, mas parte de um processo psicológico que origina e mantém todos os restantes comportamentos de exposição.

No passado, este factor foi associado a ambas as principais perspectivas explicativas da exposição televisiva. Porém, a investigação mais recente associa-o em maior grau às teorias estruturais que defendem uma menor intencionalidade subjacente às decisões de acesso ao *media* e de selecção programática.

Quem defende uma concepção mais passiva do comportamento da audiência associa o hábito às teorias que veem o *media* como uma mera actividade lúdica, que permite preencher os tempos livres e os tempos que intercalam actividades de trabalho e de lazer. Consequentemente, neste último contexto, a disponibilidade em vez da selectividade desempenha um maior poder explicativo da exposição.

III.4.4. Os modelos de *Ratings* e as perspectivas explicativas da exposição

A previsão das audiências é essencial para a sobrevivência das organizações, mas, infelizmente, atendendo às dificuldades que a indução coloca nas Ciências Sociais, os resultados alcançados não têm sido satisfatórios. A inexistência de uma perspectiva integrada explicativa dos comportamentos também contribui para esta tendência.

➤ A abordagem das gratificações e usos do *media* defende o carácter activo dos telespectadores sendo as suas escolhas função de preferências ou da satisfação de determinadas necessidades. A televisão neste contexto será um veículo de entretenimento e de informação, entre outras possíveis motivações, existindo uma selecção maioritariamente activa e intencional (*Ronsenbegren et al., 1985*).

➤ A teoria dos factores estruturais explica melhor o comportamento da audiência quando assistem com passividade, aderindo às condicionantes estruturais e recorrendo à televisão com maior ou menor regularidade para preencher os seus tempos livres ou vazios. Sob esta perspectiva será uma utilização ritualizada condicionada pela disponibilidade da audiência e influenciada por factores estruturais, entre os quais, os comportamentos estruturais associados à duplicação das audiências. Os “modelos de *ratings*” podem ser essencialmente enquadrados nas abordagens estruturais de análise do consumo televisivo.

Os modelos explicativos e de previsão desenvolvidos segundo as teorias estruturais têm permitido alcançar os melhores resultados, e não os que se baseiam na análise dos conteúdos programáticos e das necessidades ou interesses dos telespectadores. As sondagens de intenção de ver apresentam margens de erro que podem alcançar os 200%.

Os modelos explicativos de *ratings* ao nível dos indicadores agregados procuram, essencialmente, os factores estruturais que determinam a exposição das audiências (Cooper, 1993; Gensh & Sharman, 1980; Henry & Rinne, 1984). Os factores estruturais podem-se repartir por duas categorias: (a) as características de longo prazo do mercado e fora do controle das estações televisivas, por exemplo, o nível de penetração da TV_Cabo; e (b) os factores temporais e de planeamento, como, por exemplo, as variáveis de desempenho e caracterização de programas, canais, espaços e inserções comerciais.

As análises matemáticas de previsão ou explicativas podem ter em conta estes factores, cada um com o seu peso correspondente. Em virtude da existência de erros de amostragem e de uma certa variabilidade associada a cada factor se os modelos são muito ambiciosos, inevitavelmente, a precisão dos resultados proporcionados será muito reduzida. Os modelos devem seleccionar apenas os factores mais relevantes e promover a identificação de relações directas e causais entre as variáveis. Contudo, os modelos causais podem eventualmente não ser viáveis nesta área. Os autores Wulf, Verhagen e Warty (2006) apresentam uma síntese ao nível das metodologias de investigação adoptadas e desenvolvem um sistema que visa a geração de padrões relevantes e contextuais que possibilitem a previsão de comportamentos de consumo televisivo.

De acordo com Weber (2002), na bibliografia da especialidade existe um forte deficit ao nível das análises explicativas e de previsão das audiências recorrendo a abordagens não lineares e híbridas e a metodologias mais recentes de análise de dados. "Many previous forecasting models have used Regression or a logit model for channel share, with no models using the state-of-the-art methods available today." (Danaher et al., 2011, p.8).

Os modelos de *ratings* condicionam a estruturação da grelha e a venda do espaço comercial. Devemos tentar perceber o que condiciona a evolução do comportamento agregado. Os estudos desenvolvidos no âmbito dos modelos de *ratings* pretendem identificar os factores estruturais, e não individuais associados à perspectiva de gratificações e usos do *media*, que determinam as audiências (Cooper, 1993; Gensh & Sharman, 1980; Henry & Rinne, 1984).

III.5. A INVESTIGAÇÃO REALIZADA NO SECTOR TELEVISIVO E ENQUADRAMENTO DA TEMÁTICA PROPOSTA

III.5.1. As temáticas em análise no sector TV

No sector televisivo facilmente identificamos uma grande diversidade de núcleos temáticos, disciplinas, instrumentos, orientações e áreas de investigação. Na seguinte *Figura III.5.1.1* constam algumas das disciplinas e áreas de análise que podem ser identificadas na bibliografia.

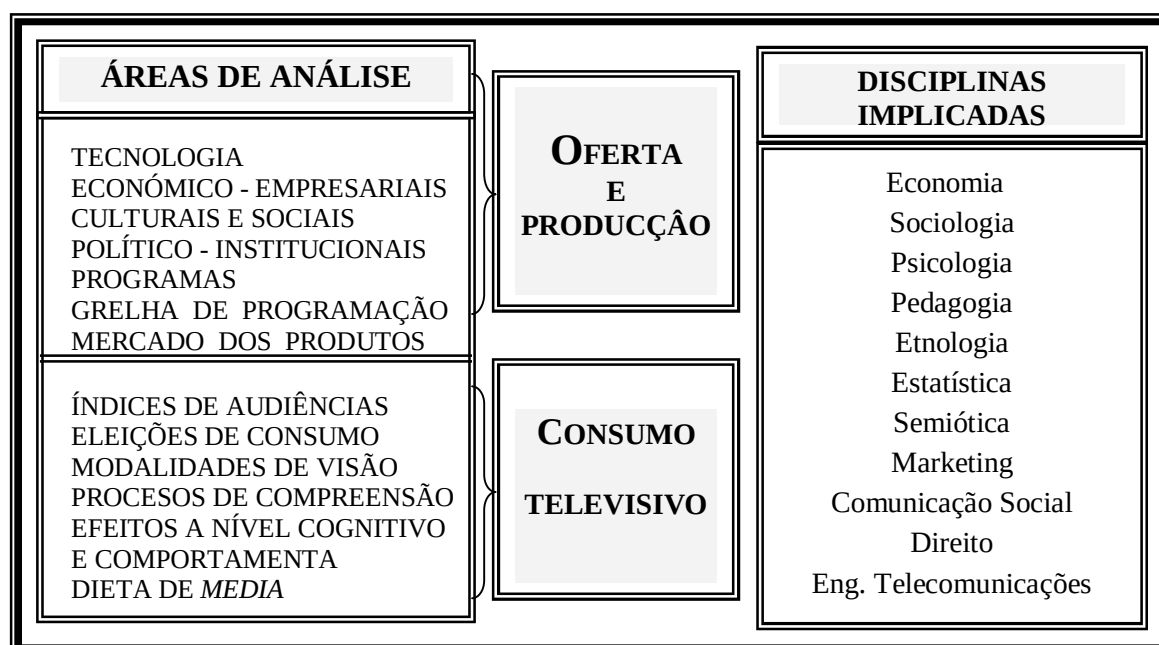


Fig. III.5.1.1 –Âmbitos de investigação na área televisiva e disciplinas implicadas.

Fonte: Casseti & Chio (1997).

Alguns trabalhos abordam questões económico-empresariais e a análise dos índices de audiências pode ser enquadrada na análise do consumo televisivo. A audimetria é uma das fontes de dados adoptada na investigação académica e empresarial e viabiliza a construção dos modelos de *media* e de alguns modelos explicativos do consumo televisivo. Os investigadores, essencialmente, anglo-saxónicos recorrem a esta fonte de dados quando o objectivo das análises reside em contribuir para as necessidades de informação no decorrer das diversas etapas que integram a construção de uma campanha publicitária. As principais disciplinas implicadas na sua análise são a Economia, o Marketing e as Ciências Matemáticas.

De acordo com *Stavitsky (2000)*, os dados designados por “*ratings*” viabilizam inúmeras análises quanto ao impacto social e económico deste meio audiovisual. Os “*ratings*” medem a exposição a programas, espaços comerciais e canais e apresentam inúmeras vantagens face aos métodos complementares de medição de audiências, nomeadamente o seu maior rigor, rapidez e a grande diversidade de indicadores que caracterizam a exposição para reduzidos intervalos de tempo e de forma diacrónica. A caracterização e a previsão dos comportamentos depende da sua quantificação para um grande horizonte temporal e para uma elevada dimensão da amostra, o que apenas será viável recorrendo aos painéis e à audimetria. Outras fontes de dados podem complementar as análises, mas os sistemas de informação são essenciais para a construção dos modelos de *media*.

III.5.2. A Investigação Académica e Empresarial

No passado, tem existido um grande distanciamento entre os estudos comerciais e os académicos, sendo, frequentemente, considerados como caminhos opostos e apresentando objectivos e interesses que não são conciliáveis. De acordo com *Huertas (2002)*, cada vez mais será imprescindível sublinhar o contributo de um enfoque multidisciplinar.

A autora refere que após décadas de esforços no sentido de encontrar soluções de um ou do outro lado, a disjunção entre um conhecimento académico e um conhecimento “administrado” das audiências encontra-se neste momento mais aberto do que nunca.

O objectivo proposto neste trabalho incide nas necessidades de informação do meio empresarial, mas avaliando o contributo de uma abordagem multidisciplinar para que sejam alcançados melhores resultados quando se pretende a construção dos modelos de exposição.

III.5.3. As metodologias aplicadas aos estudos de audiências

A existência de duas perspectivas antagónicas associadas à análise da exposição televisiva resulta do conjunto muito abrangente de disciplinas que concorrem para a sua análise, e que nomeadamente se pode classificar em dois grandes segmentos: a investigação qualitativa e a quantitativa das audiências televisivas. Os métodos qualitativos estão orientados para o conhecimento e compreensão de atitudes, opiniões e motivações e os métodos quantitativos são aplicados com o intuito de conhecer um tipo de informação quantitativa em forma de estatística (*Espinosa, 2006*)

Verificamos que esta divisão de perspectivas, bem notória na bibliografia da especialidade, não tem contribuído para que neste sector se promova um maior conhecimento sobre como a população interage com o televisor. A oposição entre o qualitativo e quantitativo corresponde a uma oposição clássica da tradição epistemológica, que indica também divisões académicas e distinções no tipo de organizações (*Espinosa, 2006*).

A maioria dos trabalhos quantitativos publicados no âmbito da audimetria, nas principais revistas da especialidade, referem-se a análises realizadas em universidades Norte Americanas, Britânicas ou em outros países anglo-saxónicos e em muito menor grau nos restantes países Europeus.

Técnicas Qualitativas	Técnicas Quantitativas
Motivos, atitudes, comportamentos, formulação de hipóteses e avaliação das mesmas	Quantificação de dados e extrapolação dos mesmos para inferir resultados
Análises não Estatísticas	Análises Estatísticas
Amostras geralmente reduzidas	Amostras de maior dimensão
Recolha de informação não estruturada	Recolha de informação estruturada

Quadro III.5.3.1 – Diferenciação entre as abordagens qualitativas e quantitativas.
Fonte: Espinosa (2006)

Os escassos trabalhos europeus e não anglo-saxónicos identificados, perante esta lacuna bibliográfica, por vezes, consistem em estudos comparativos face ao que se verifica, por exemplo, nos Estados Unidos. Atendendo ao reduzido número de trabalhos nacionais relacionados com os modelos de “*ratings*”, a presente tese visa contribuir para um maior conhecimento sobre a evolução dos dados de audimetria.

Em *Huertas (2002)*, refere-se a existência de divergências quanto aos métodos de investigação na área da comunicação de massas entre os dois continentes. Segundo a autora, quando se classifica de forma global as abordagens existentes nesta área, devemos recorrer ao seu lugar de procedência, distinguindo-se entre os Estados Unidos e a Europa. As diferenças contextuais, históricas, económicas, políticas e sociais entre ambos os continentes contribuíram para que o estudo da comunicação social de massas seja diferente. Salienta que foram estabelecidas directrizes divergentes e inclusive contraditórias em algumas ocasiões, determinadas pelo país de origem, devido à diversidade de teorias e das ferramentas de análise. Também, é referido o confronto que se verifica entre o quantitativo, mais estendido na América do Norte e o qualitativo que tem uma maior implantação na Europa e na América Latina, e que exemplifica esta questão.

No que diz respeito ao nível da maior ou menor adequabilidade das abordagens metodológicas quantitativas versus qualitativas para a análise do consumo televisivo em *Huertas (1997, 2002)*, *Souchon (1993)*, *Saperas (1998)* e *Igartua e Badillo (2003)* podemos encontrar uma excelente exposição sobre esta temática. Para *Huertas (2002)* é conveniente ressaltar que não existe objectividade, nem neutralidade quando se trata do estudo de aspectos sociais e culturais. A nossa visão dos factos dependerá do método utilizado pelo investigador. Salienta que não existe uma visão mais correcta e que a prova está em que nenhuma prevaleceu sobre as restantes. A escolha da abordagem depende de quem realiza a investigação e com que finalidade. Um investigador, com uns princípios e seguindo determinados passos metodológicos, obterá uma informação que complementar, negará ou modificará as visões de outros analistas com bases e caminhos metodológicos diferentes. Só, assim, será possível um maior progresso do conhecimento.

Nesse sentido, na presente tese, apresentam-se os potenciais de aplicação de áreas metodológicas mais recentes, não só para o desenvolvimento de modelos de “*ratings*”, mas também para a análise da interacção dos telespectadores com o meio televisivo. O quantitativo contribui com uns dados finais e o qualitativo oferece um tipo de informação que ajuda a compreender o estudado ao contribuir com argumentos que explicam o que no primeiro caso aparece reflectido por valores mensuráveis, absolutos ou percentuais (*Huertas, 2002*).

Por conseguinte, as abordagens metodológicas que visam extrair o máximo de informação das bases de dados de audimetria contribuem para a identificação de situações, factos, comportamentos que, posteriormente, podem ser desenvolvidos pelas áreas que têm essa responsabilidade. Nomeadamente, também existe a necessidade de combinar as possibilidades do quantitativo e do qualitativo, uma vez que a melhor via para a análise deste sector, no futuro, será a da análise interdisciplinar (*Huertas, 2002*). Cada parceiro da investigação deve contribuir com a aplicação dos seus métodos, disponibilizando aos restantes os seus resultados, ideias e comentários, numa tentativa combinada de conseguir extrair um maior conhecimento sobre os fenómenos em análise (*Fonseca, 2008*).

Na perspectiva da construção dos modelos de exposição, ambas as perspectivas devem ser consideradas. Os modelos quantitativos permitem a análise da estrutura dos comportamentos de consumo televisivo e da sua evolução, enquanto que o qualitativo em determinadas etapas permitirá uma melhor definição das variáveis a considerar nas análises que visam a extracção de regularidades estatísticas e que, por sua vez, viabilizam a previsão da exposição.

III.5.4. Estado da Arte dos Modelos de “ratings”

No prefácio do livro intitulado “*Ratings Analysis: Theory and Practice*” (Webster & Litchy, 1991), os seus autores apelam à utilização destas bases de dados na investigação académica. No entanto, em virtude da dificuldade de acesso aos dados e na respectiva análise, assim como pelo valor comercial que as análises apresentam verificamos que existem escassas publicações internacionais acessíveis aos investigadores académicos (Danaher et al., 2011; Gensh & Shaman, 1980; Weber, 2002) e em especial a nível nacional onde apenas foi possível identificar um número muito restrito de trabalhos.

Segundo Stavitsky (2000), os trabalhos publicados que visam a análise de “ratings” podem-se agrupar, essencialmente, em três áreas de análise:

➤ A análise das estratégias de programação e de contra programação e a identificação de programas com maiores níveis de audiência para serem objecto de análises posteriores (Eastman & Otteson, 1994; Eastman, Newton & Gregory, 1998; Eastman, Newton & Pack, 1996; Henry & Rinne, 1984).

Neste item inclui-se a análise da repercussão de determinadas estratégias de contra programação, através da transmissão de tipologias consideradas como complementares face às da concorrência; a avaliação dos efeitos decorrentes de estratégias de “Lead_In” e de “Lead_Out” que consistem na transmissão de um programa mais popular que beneficia as audiências do programa posterior ou anterior respectivamente; a análise do

“Blunting”, ou seja, da transmissão de um programa que coincide ao nível da tipologia face aos da programação dos canais concorrentes, mas que apresenta uma maior exposição; a análise do planeamento defensivo transmitindo-se o “melhor” programa de um canal em simultâneo com o programa mais popular dos canais concorrentes e a análise do “planeamento ofensivo” quando se transmite um programa com os maiores níveis de exposição na mesma faixa horária em que são transmitidos os programas mais fracos dos canais concorrentes.

- A caracterização de inúmeros aspectos relacionados com a exposição televisiva, considerando a vertente programática e a comercial, como, por exemplo, a análise do “zapping” que consiste na transição dos telespectadores entre canais, nomeadamente, quando são transmitidos os espaços comerciais e a análise da repercussão das promoções de programas no seu futuro desempenho (*Alwitt & Prabhaker, 1992; Burnett, 1991; Davis & Westbrook, 1985; Gantz, 1985; Harwood, 1997; Rosenstein & Grant, 1997*).
- A análise das forças estruturais da indústria, como, por exemplo, os trabalhos que visam a análise dos índices de duplicação e os modelos de previsão e explicativos de “ratings”, enquadrados no âmbito das teorias estruturais explicativas da exposição televisiva.

Na realidade os trabalhos de previsão são muito escassos, sendo este aspecto salientado pelos autores dos trabalhos identificados nesta área, como consequência das dificuldades que o acesso aos dados coloca e que as análises de indução apresentam nas Ciências Sociais e mais concretamente na área da audimetria. Do ponto de vista explicativo, também, são escassas algumas vertentes de análise, como, por exemplo, em relação aos factores que condicionam a evolução dos principais indicadores de audimetria em função dos quais se comercializam os suportes publicitários.

Apenas foi identificado na bibliografia da especialidade, um número muito restrito de modelos explicativos e de previsão quantitativos de valores de “ratings”, essencialmente, internacionais recorrendo, maioritariamente, a metodologias lineares e,

muito pontualmente, às abordagens não lineares (Garland, 1996, Geel, Hoogstrate & Foley, 2003; Gensh & Shaman, 1980; Malara, Armellini & Monotti, 2004; Xiaoling & Hing-Po Lo, 2006; Weber, 2002; Wulf et al., 2006). Verifica-se a mesma tendência metodológica quando se abordam as diversas perspectivas de análise dos comportamentos de duplicação das audiências onde os autores frequentemente, também, recorrem a variáveis pouco adequadas às análises propostas. Também, são escassos os trabalhos que promovem a avaliação da repercussão da estruturação do mercado televisivo, como, por exemplo, o grau de penetração da TV Cabo e dos factores de planeamento (Cooper, 1993) na evolução dos “ratings”.

EFEITOS DE HERANÇA ENTRE PROGRAMAS ADJACENTES	LEALDADE TIPOLOGICA E AO CANAL
EXPOSIÇÕES REPETIDAS	ANÁLISES DE PREVISÃO E EXPLICATIVAS DE INDICES DE AUDIÊNCIAS
ESTRUTURAÇÃO DO MERCADO	REPETIÇÃO DE SELECÇÕES
ESTRATÉGIAS DE PROGRAMAÇÃO E DE CONTRA-PROGRAMAÇÃO	FACTORES DE PLANEAMENTO A NIVEL COMERCIAL
	FACTORES DE ACESSO E DISPONIBILIDADE

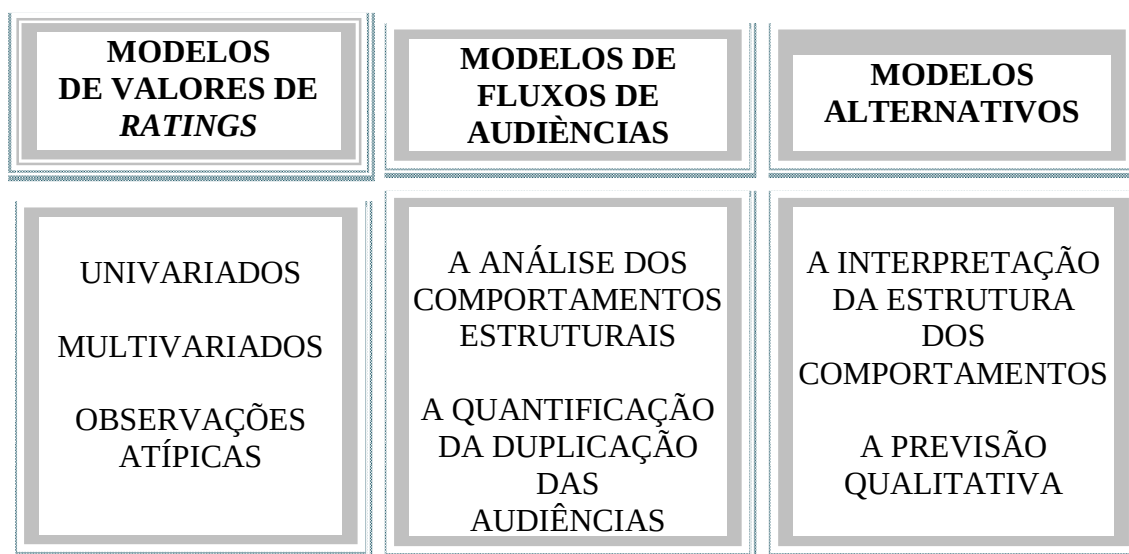
Alguns dos trabalhos mais recentes que foi possível identificar na área da previsão correspondem a duas teses de Doutoramento internacionais, uma das quais realizada no Reino Unido, e que consistiu na previsão das audiências dos jogos de futebol, recorrendo a abordagens qualitativas (Buraimo, 2007). Xiaoling (2007) procurou dar sugestões quanto à estruturação da grelha de programas por aplicação de metodologias de Estatística Bayesiana a bases de dados de audimetria chinesas e na área dos índices de duplicação realizou algumas análises aos índices que quantificam a lealdade programática (Xiaoling & Hing-Po Lo, 2007). Em Patelis et al. (2003) foi realizada a previsão de índices de audiências gregos, recorrendo aos valores das suas séries temporais e a modelos de resposta do mercado.

Fig. III.5.4.1 – Âmbitos de análise segundo a perspectiva estrutural de análise da exposição televisiva. Fonte: Elaborado pela autora.

Para resumir o estado da arte vamos atribuir uma maior relevância aos trabalhos realizados na área da análise dos factores estruturais, que visam o desenvolvimento de modelos explicativos e/ou de previsão de audiências programáticas e comerciais. A *Figura III.5.4.1* ilustra os campos de aplicação das teorias estruturais da exposição televisiva e que abrange, nomeadamente, as análises explicativas ou de previsão de “*ratings*”, enquanto tipologia de dados associada à tecnologia de audimetria de medição das audiências televisivas. Na vertente explicativa identificamos as análises de duplicação das audiências, a análise das estratégias de programação e de contra-programação e a análise de determinados comportamentos de consumo televisivo.

III.6. ESTADO DA ARTE DOS MODELOS QUANTITATIVOS DE EXPOSIÇÃO TELEVISIVA

Uma vez que a bibliografia da especialidade incide, essencialmente, nos modelos de *ratings*, inicialmente, apresenta-se uma síntese das diversas perspectivas que foram identificadas, ao nível dos modelos explicativos e/ou de previsão dos principais indicadores de audimetria, que condicionam a venda da publicidade televisiva. Na sua sequência referem-se os modelos alternativos de construção dos modelos de exposição.



Quadro III.6.1 – Enquadramento do estado da arte na área dos modelos de exposição.

Fonte: Elaborado pela autora.

As áreas que condicionam maioritariamente a actividade de planeamento de *media* correspondem, essencialmente, às seguintes:

(1) A previsão dos valores de *ratings* referentes a programas, suportes horários, inserções comerciais e dos factores bloco que quantificam a queda das audiências no decorrer dos espaços comerciais. Segundo *Danaher et al. (2011)*, correspondem a áreas que têm sido negligenciadas na investigação académica internacional e a nível nacional apenas foi possível identificar alguns trabalhos que incidem nos valores dos *shares* de audiências (*Datia, 2002*);

(2) Os modelos explicativos da evolução dos principais indicadores de audimetria, nomeadamente dos valores de *ratings*, das taxas médias de audiências para o total da TV e dos *shares* de audiências. As análises explicativas diferenciam-se das análises de previsão, uma vez que no contexto multivariado podem incluir factores que não são conhecidos ou previsíveis no momento de realizar uma previsão. Por outro lado, uma maior percepção sobre os factores que condicionam a evolução dos indicadores, também, é relevante, uma vez que actuando sobre essas dimensões podemos alterar a sua evolução e maximizar a exposição ao mercado alvo, assim como contribuir para uma maior precisão das previsões que forem viáveis e mais adequadas à evolução dos dados;

(3) A previsão quantitativa da exposição de um colectivo face a um conjunto de suportes. Apenas foi viável identificar em *Zarate (1997)* uma sequência metodológica que permite a identificação dos suportes a incluir em planos de *media* considerando a exposição colectiva e determinadas metodologias de estatística multivariada.

(4) As teorias explicativas da exposição, em especial os modelos enquadrados na perspectiva estrutural e na perspectiva das gratificações e usos do *media*. Os modelos existentes são na sua maioria parciais e não integrativos das teorias existentes;

(5) As análises de duplicação das audiências enquadradas nas teorias estruturais explicativas da exposição televisiva. Os indicadores acumulados permitem a análise dos “comportamentos estruturais” e subdividem-se em análises de lealdade programática, lealdade tipológica e ao canal, efeitos de herança e exposições repetidas. Contribuem para a avaliação da regularidade da exposição a nível agregado, que é essencial para as análises de previsão, assim como para a construção de um modelo estrutural explicativo dos padrões de consumo televisivo ou integrado que permite avaliar a viabilidade da previsão dos fluxos de audiências entre os suportes. Os indicadores acumulados de audimetria permitem a interpretação da evolução dos *ratings* e avaliação da sua “qualidade”, uma vez que as formas de consumo mais dinâmicas sobrestimam os valores das audiências.

III.6.1. Os modelos de exposição ao nível dos indicadores agregados ou “gross meares” de audimetria

Quando se pretende construir um plano de *media* poderá ser relevante a previsão dos valores dos *ratings* dos suportes considerados e uma maior percepção sobre os factores que condicionam a sua evolução. No *Quadro III.6.1.1* constam os trabalhos publicados que visam a sua previsão.

De acordo com *Danaher et al. (2011)*, apenas conseguiram identificar dois trabalhos publicados de previsão de valores de *ratings*, considerando bases de dados a partir do ano 2000. As análises incidem em análises de Regressão linear ou Logit, recorrendo a dados agregados das faixas horárias da noite e prevalecendo como método de recolha de dados os diários de audiências. A bibliografia é essencialmente anglo-saxónica. Ao nível dos restantes indicadores de audimetria verifica-se uma maior lacuna na bibliografia disponível e as mesmas limitações ao nível das metodologias e fontes de dados utilizadas.

Autor	Variável	País	Fonte de Dados	Metodologias
Darmon (1976)	<i>Ratings</i>	Canada	Diários	Regressão
Danaher & Mawhitney (2001)	<i>Ratings</i>	Nova Zelândia	Diários	Logit model
Kelton & Stone (1998)	<i>Ratings</i>	US	Diários	Regressão
Reddy, Aronson & Stam (1998)	<i>Ratings</i>	US	n/s	Regressão
Rust & Alpert (1984)	<i>Ratings</i>	US	Inquérito	Logit model
Weber (2002)	<i>Ratings</i>	Alemanha	Audimetria	Rede Neuronal

Quadro III.6.1.1 - Trabalhos que incidem na previsão dos valores de *ratings*.

Fonte: Exerto do estado da arte apresentado em *Danaher et al. (2011)* a respeito dos diversos indicadores de audimetria.

III.6.1.1. Os modelos de exposição univariados

Nos modelos univariados não se considerarem variáveis exógenas, apenas desfasamentos da variável a prever e são na sua maioria modelos sazonais ou de distribuição (*Weber, 2002*). Em particular estes modelos recorrem ao ajustamento de funções de densidade acumulada às médias observadas da variável em estudo. *Goodhart, Ehrenberg e Collins (1975)* foram alguns dos autores pioneiros nesta área e tentaram prever a audiência acumulada de episódios subsequentes de um programa. Recorrendo a uma distribuição empírica dos telespectadores que assistiram aos primeiros pares de episódios de uma série programática ajustaram uma função de densidade acumulada teórica e de preferência a distribuição Beta, assim como em *Headen, Klompmaker, Teel e Jessie (1977)*.

Os modelos sazonais recorrem a dados de séries temporais e requerem séries equitativamente espaçadas no decorrer do tempo, utilizando-se em geral períodos pré-definidos.

Estes modelos calculam as previsões com base na análise da sazonalidade do consumo histórico. Assim, os modelos distribucionais, negligenciam variáveis que podem ser relevantes para a previsão e para a compreensão dos resultados proporcionados. Ainda que os dados requeridos para a sua construção sejam simples e estando as metodologias de análise de séries temporais (ST) implementadas na maioria dos *softwares* de análise estatística, verificamos que existem poucas publicações que demonstrem a adequabilidade de metodologias mais recentes a este nível.

Um dos primeiros modelos sazonais foi desenvolvidos por *Gensch e Shaman (1980)*. Os autores realizam análises de previsão de valores de taxas médias de audiências para o total da TV considerando suportes do horário nobre que correspondem a intervalos de 15 minutos e utilizando como abordagem metodológica a Regressão Trigonométrica e a Análise de Fourier. Os métodos clássicos de previsão univariada de audiências, como, por exemplo, as análises ARIMA ou as análises de Fourier apresentam como vantagem a sua simplicidade e a disponibilidade de métodos para controlo do sobre ajustamento aos dados.

Uma vez que estas abordagens negligenciam variáveis que podem ser relevantes para as análises de previsão, não será viável uma interpretação dos resultados alcançados. Outro aspecto negativo a salientar corresponde à necessidade de requerem observações regularmente espaçadas o que acaba por gerar amostras aleatórias demasiado grandes para conseguir abranger um programa por completo. A sua utilização é restrita, atendendo às desvantagens que apresentam na prática, e apenas se adequa a determinados contextos onde se possa observar uma maior regularidade na evolução dos indicadores. Na opinião de alguns autores (*Weber, 2002*), do ponto de vista clássico será mais recomendável a aplicação de modelos GLM, uma vez que permitem uma interpretação dos resultados alcançados através das variáveis identificadas como sendo mais relevantes.

Os modelos sazonais não lineares recorrendo a abordagens da área da Inteligência Artificial, também são praticamente inexistentes.

Uma realidade que é considerada surpreendente, por alguns autores (*Weber, 2002*), uma vez que, frequentemente, surgem em outras áreas de complexidade comparável, como, por exemplo, na área da previsão das séries económicas ou de outros dados financeiros. *Garland (1996)* realizou algumas análises com dados de audimetria Australianos, recorrendo às Redes Neurais para tentar prever a evolução das séries temporais de valores de *ratings*. Utiliza dados de *ratings* ao nível do quarto de hora, desde a sexta até à novena semana de 1995. Não foi possível identificar publicações que utilizem metodologias mais recentes de previsão não linear de séries temporais univariados.

III.6.1.2. Os modelos de investigação multivariados programáticos

Os modelos explicativos e/ou de previsão multivariados podem ser lineares ou não lineares. Alguns modelos explicativos lineares, utilizando combinações de valores históricos médios agregados encontram-se descritos em *Cattin, Festa e Le Diberder (1994)*. *Darmon (1976)* desenvolveu um modelo de Regressão linear para avaliar a influência da tipologia programática e da lealdade ao canal ao nível da selecção dos programas. *Zufreden (1973)* considerou o conceito de inércia como um factor determinante do comportamento de consumo televisivo.

Na área do planeamento comercial, no final dos anos oitenta, foram desenvolvidos alguns modelos lineares para tentar prever os valores das audiências comerciais (*Cooper, 1993; Gensch & Shaman, 1980; Goodhart et al., 1975*). Estes modelos avaliam a repercussão das variáveis temporais, da sazonalidade do consumo televisivo, do conteúdo programático e por vezes, de um índice que pretende quantificar a atractividade que os telespectadores manifestam por determinados programas (*Bortz, 1986*). *Goodhart et al. (1975)* descobriram que os níveis de audiência dos programas adjacentes influenciam o desempenho de um programa, designando estes efeitos de “Lead_in” e de “Lead_out”.

Gensch e Shaman (1980) desenvolveram um modelo de Regressão Trigonométrica aplicado a dados de séries temporais para tentar prever os valores de *ratings* de novas combinações de programas, considerando suportes previamente existentes e novos conteúdos. *Henry e Rinne (1984)* desenvolveram um modelo baseado na metodologia Logit que pretende identificar quais são as variáveis que influenciam a evolução dos *ratings* programáticos.

O consumo televisivo, também, pode ser modelado de forma não linear, mas os trabalhos publicados são ainda mais escassos. Apenas foi identificado um número muito reduzido de trabalhos que utilizam metodologias da área da Inteligência Artificial para realizar as previsões.

Segundo *Weber (2002)*, os trabalhos existentes recorrem às Redes Neurais para prever os valores de *shares* de audiências e os denominados Sistemas Expertos para realizar uma previsão não linear, como, por exemplo, em *Fitzsimons, Khabaza e Shearer (1993)*. Os autores, em geral, consideram como variáveis explicativas e de previsão: a sazonalidade, a classificação do conteúdo na forma de tipologias programáticas, os conteúdos em contra- programação e a envolvente do programa. De acordo com *Weber (2002)*, não são referidos, pormenores a respeito da definição da contra-programação e da classificação dos conteúdos, assim como, também, não constam especificações dos períodos a prever e do processo de estimação.

Nikolopoulos, Goodwin, Patelis e Assinkopoulo (2006) utilizaram a Regressão linear Múltipla, as Redes Neurais e algoritmos *Instance Based Learners* (IBL) para avaliar a repercussão da transmissão dos jogos de futebol nos valores dos *shares* recorrendo a dados de audimetria gregos. Segundo *Danaher et al. (2011, p.2)*: “A shortcoming in earlier works includes a tendency not to validate forecasting models or doing so for a small number of programs.” Os autores descrevem uma aplicação da análise bayesiana das séries temporais multivariadas de valores de *ratings*. Em *Danaher e Dagger (2012)* consta uma proposta de análise “nested logit” para prever as decisões de selecção programática do ponto de vista agregado.

Em *Assinkopoulo (2010)* avalia-se a capacidade predictiva da RL, algoritmos *K Nearest Neighbours* (KNN) e das Redes Neurais na área das audiências dos jogos de futebol que geram observações atípicas.

Os modelos híbridos combinam várias abordagens (*Weber, 2002*). Uma vez que o seu desempenho é superior aos dos modelos uni-métodos, estas abordagens são referenciadas de forma crescente na bibliografia da teoria econométrica. No entanto, não foi possível identificar esta tipologia de modelos em trabalhos publicados sobre a previsão das audiências televisivas. Em suma, verificamos que existe um grande deficit ao nível dos modelos explicativos e de previsão não lineares e híbridos das audiências televisivas, assim como referentes à análise das séries temporais (*Weber, 2002*).

Apenas foi possível identificar dois trabalhos que analisam a dimensão dos erros de previsão das audiências televisivas, e que constam em *Yoo e Kim (2002)* e em *Napoli (2001)*. Em *Yoo e Kim (2002)* descrevem-se análises de Regressão linear Múltipla que avaliam a repercussão de alguns factores estruturais na dimensão dos erros de previsão. As análises incidem em bases de dados dos Estados Unidos da América (USA) e da Coreia. Os factores em análise correspondem a variáveis temporais, como, por exemplo, se o programa é transmitido no decorrer da semana ou ao fim de semana, as tipologias programáticas dos programas adjacentes e do programa em análise, assim como a duração do programa e dos programas de tipologias equivalentes transmitidos em contra programação.

Os modelos de Regressão linear Múltipla apresentam uma capacidade explicativa de apenas 24% se consideramos os dados Coreanos e de 19% para os dados dos USA, onde apenas é relevante o dia e a tipologia do programa antecessor. A dimensão dos erros é menor ao fim-de-semana quando se mantêm as tipologias adjacentes, assim como para os programas de comédia. Com dados de audimetria da Coreia, apenas são significativas as variáveis que representam a duração dos programas e a tipologia programática de drama.

Os autores salientam, que atendendo aos resultados alcançados, existe uma forte necessidade de considerar outros factores, tais como a popularidade dos programas adjacentes, o canal, o grau de penetração da TV Cabo, entre outras variáveis.

Relativamente à análise do efeito do nível de agregação dos dados na precisão das previsões de valores de *ratings* apenas foi possível identificar o trabalho de *Meyer e Hyman (2005)* onde se conclui que as análises de previsão devem ser realizadas para segmentos ou para o universo televisivo. Na opinião dos autores, as análises realizadas a nível individual produzem maiores valores dos erros devido às particularidades do seu comportamento. Os segmentos a considerar devem ser os mais homogéneos em relação aos padrões de exposição televisiva dos telespectadores que os constituem e não devem ser construídos apenas com base nas variáveis sociodemográficas.

III.6.1.3. Os modelos de exposição comercial

Apesar de no passado, se ter questionado a eficácia da publicidade televisiva, hoje em dia, a televisão mantém-se como o *media* mais utilizado para a divulgação de produtos e serviços. A contínua subida dos custos publicitários e a fragmentação das audiências tem suscitado um crescente interesse dos profissionais do meio por inúmeras temáticas relacionadas com a exposição aos espaços comerciais. Na área comercial existem, essencialmente, duas linhas de investigação: a presencial e os estudos de “recall”, ou seja, a análise das taxas de recordação das inserções publicitárias. Os trabalhos desenvolvidos ao nível das taxas de recordação procuram avaliar a repercussão da transmissão dos espaços comerciais nas decisões de compra dos telespectadores e contribuem para o desenvolvimento de estratégias de planeamento comercial que maximizem a exposição e, consequentemente, também a memorização dos anúncios publicitários. A investigação presencial aborda a questão da exposição aos espaços comerciais e neste contexto identificamos as seguintes áreas: “Zapping” e “Zipping” de espaços comerciais; a influência da duração e da posição dos anúncios; os efeitos da programação na publicidade televisiva e os espaços comerciais e o “Clutter”, entre outras.

Alguns trabalhos têm demonstrado que as taxas de recordação são reduzidas, inclusive quando se considera um curto espaço de tempo após a exposição. O público com uma menor tendência para evitar os espaços comerciais, obviamente, recorda um maior número de marcas anunciadas, sendo a investigação presencial determinante neste contexto. Para os “zappers” as marcas posicionadas próximo do final do espaço comercial são as que maximizam esta variável e, conseqüentemente as inserções publicitárias são tarifadas de forma diferenciada em função da sua localização (*Agee, 1992; Clancy & Kweskin, 1971; Cobb, 1985; Murphy, Cunningham & Wilcox, 1979*). *Greene (2007)* afirma que estes comportamentos, nem sempre, afectam negativamente a taxa de recordação das inserções publicitárias, uma vez que para mudar de canal é necessário estar atento ao televisor. Apesar das reduzidas taxas de recordação dos anúncios, a eficácia da publicidade televisiva encontra-se certamente condicionada pelo número de pessoas que lhe estão expostas.

Quando se constrói um plano de *media* recorre-se a determinadas técnicas, como, por exemplo, ao “roadblocking” em que se transmite um anúncio simultaneamente em todos os canais considerados como mais importantes para alcançar o mercado alvo; ao “comercial wraparound” dividindo um anúncio de 30 segundos em dois de 15 segundos e a outras estratégias destinadas a maximizar a exposição e a cativar o público.

A atractividade exercida por um programa pode influenciar as taxas de recordação das diversas posições das inserções publicitárias. Se um programa não for considerado interessante pelos telespectadores, as posições finais de um intervalo podem não proporcionar o efeito desejado. Outra variável a ter em atenção será à localização do intervalo, sendo as taxas de recordação superiores para os anúncios localizados no decorrer de um programa, mas também são função do seu desempenho. A posição de uma inserção, também, condiciona a eficácia de uma inserção publicitária.

A bibliografia presencial tem demonstrado que prevalece a fuga aos anúncios através da mudança de canal no seu decorrer ou via “zipping” que caracteriza o comportamento de “fast forwarding” quando a audiência assiste a programas que foram gravados (*Heeter &*

Greenberg, 1988; Kaplan, 1985). O excessivo “clutter”, ou seja, o número de anúncios por intervalo é considerado como um dos factores que contribui para esta tendência, reduzindo a eficácia dos anúncios (*Hammer, Riebe & Kennedy, 2009*). A capacidade de um anúncio para conseguir transmitir informação considerada útil e nova pelos telespectadores influencia a intensidade do “zapping” (*Stewart, 1989*).

Pechman e Stewart (1988), Shumann, Richard, Scott et al. (1990) e Goldberg et al. (1993) analisam os efeitos que resultam de uma transmissão repetitiva dos anúncios. *Siddarth e Chattopadhyay (1998)* identificaram uma relação em forma de U ou J entre as múltiplas exposições a um anúncio e a probabilidade de se serem evitados os anúncios. A elasticidade ao “zapping” atinge o seu mínimo para 14 exposições, crescendo significativamente após este valor. A repercussão da duração dos anúncios foi avaliada por *Patzer (1991)*. *Kaplan (1985)* e *Heeter e Greenberg (1988)* encontraram maiores níveis de transição entre canais para anúncios transmitidos às horas ou meias horas em relação aos restantes períodos horários. Os autores analisaram os programas adjacentes a um intervalo, a contra programação e as características dos intervalos.

Finalmente, as consequências da atitude dos telespectadores face à publicidade foram avaliadas por *Mackensie, Lutz e Belch (1986)* e *Lee e Lumpkin (1995)*. *Alwitt e Prabhaker (1992)* abordam alguns dos factores que influenciam a atitude dos telespectadores face à publicidade e a predisposição para se evitarem os anúncios.

Os trabalhos que incidem na análise do comportamento da audiência no decorrer dos intervalos publicitários têm identificado que para além de uma forte tendência para o “zapping”, por vezes, o público ausenta-se nestes períodos para realizar breves actividades domésticas. Ainda que o televisor possa permanecer ligado no decorrer dos intervalos, o público poderá aproveitar estes períodos para ler ou falar, o que do ponto de vista da audimetria não será possível identificar. Os trabalhos desenvolvidos na área do “zapping” têm procurado caracterizar os segmentos mais aptos a desenvolver este tipo de comportamento, assim como o seu padrão e intensidade.

Contudo, os trabalhos que incidem na análise explicativa e /ou predictiva da evolução dos *ratings* comerciais também são muito escassos. A maioria dos modelos não inter-relaciona a componente programática com a vertente comercial. Como excepção pode ser citado *Danaher (1995)* que tentou quantificar a influência de determinadas variáveis de caracterização dos intervalos e dos programas em um índice que quantifica a redução das audiências comerciais em função dos valores programáticos. Também, é essencial o desenvolvimento de modelos explicativos nesta área para conseguir uma correcta estruturação de uma campanha publicitária cujo objectivo final será o de maximizar a sua taxa de exposição aos segmentos pretendidos.

Os trabalhos desenvolvidos na área comercial têm demonstrado que a estruturação da grelha de programas, o interesse que os programas despertam, a atitude da audiência e as variáveis de planeamento associadas aos espaços publicitários influenciam as taxas de recordação de anúncios e as vendas dos produtos anunciados. Não se deve dissociar a vertente programática e a comercial, uma vez que da sua influência recíproca resulta a dimensão dos índices de audiência.

A selecção programática e, conseqüentemente, os valores dos *ratings* comerciais, são função, entre outros factores, das preferências do público, da disponibilidade da audiência, da estruturação da grelha e da imagem do canal.

Resumindo, as análises presenciais são determinantes para o planeamento publicitário. Uma fracção destes trabalhos aborda a questão dos comportamentos de “zapping”, procurando identificar os factores que os condicionam e segmentando a audiência a este nível. Utilizam, essencialmente, dados recolhidos através de inquéritos e em muito menor grau os dados de audimetria do ponto de vista agregado. Não foi possível identificar trabalhos com dados que caracterizam a exposição comercial individual, assim como trabalhos referentes a análises de previsão. Também, são muito escassos os modelos explicativos e de previsão dos valores de *ratings* comerciais ou considerando a sua relação face aos valores programáticos.

III.6.2. Os modelos de exposição ao nível dos indicadores acumulados de audimetria: A duplicação das audiências

Os indicadores acumulados de audimetria permitem a interpretação da evolução dos indicadores agregados e quantificam as audiências conjuntas entre os suportes publicitários. Consequentemente, permitem avaliar a regularidade da exposição ao *media*, a viabilidade da sua previsão e a identificação das perspectivas metodológicas que permitem alcançar os melhores resultados. Os comportamentos que resultam da estrutura da grelha de programas designam-se de comportamentos estruturais e podem ser classificados em lealdade tipológica, ao canal, programática, efeitos de herança e exposições repetidas, sendo a sua análise viabilizada através destes indicadores.

III.6.2.1. Os efeitos de herança entre suportes

O conceito de efeitos de herança entre programas traduz os índices de duplicação das audiências entre programas adjacentes ou próximos e do ponto de vista teórico associa-se com a perspectiva que defende uma maior passividade e inércia dos telespectadores. Segundo alguns autores (*Cooper, 1993; Webster, 1985*), os programas herdam audiência dos programas que os antecedem ou sucedem, sendo a primeira tendência mais relevante. Consequentemente, incentivando o público a permanecer sintonizado no mesmo canal ao longo de sucessivos programas poder-se-á maximizar o desempenho de determinado canal (*Boemer, 1987; Cooper, 1993; Davis & Walker, 1990; Walker, 1988; Webster, 1985*).

A aderência da indústria e dos académicos ao poder dos efeitos de herança baseia-se no pressuposto implícito de que o público é de certo modo passivo, assistindo a um programa simplesmente porque sucede um programa que estão a assistir. Independentemente de isto ser verdade, os programas adjacentes por vezes partilham uma elevada percentagem da audiência e um grande número de pessoas assiste à televisão em períodos adjacentes (*Webster, 1985*), consequentemente a correlação tem que ser elevada.

O número de opções disponíveis na interface de dois programas poderá penalizar esta tendência (Davis & Walker, 1990; Tiedge & Ksobiech, 1986; Walker, 1988; Webster, 1985).

Alguns autores defendem que se o público demonstrar uma maior preferência por determinadas tipologias de programas é provável que, também, assistam ao programa subsequente se este coincidir ao nível da tipologia programática (Davis & Walker, 1990; Tiedge & Ksobiech, 1986; Webster, 1985). No entanto, existem algumas dificuldades a este nível e a primeira e mais relevante será a tipologia mais apropriada para classificar um programa. Verificamos que existe algum consenso quanto à necessidade das categorias serem definidas em função da exposição do público e evitar a utilização das tradicionais e pré-determinadas pela indústria televisiva.

Para a sua definição têm-se utilizado as análises Factoriais, mas com pouco sucesso. Porém, também é salientado que se essas categorias existirem parecem corresponder às tradicionais. Uma segunda limitação consiste no denominado efeito de saturação (Lehman, 1971), segundo o qual ainda que os telespectadores optem maioritariamente por determinados conteúdos, a partir de determinado número de programas podem perder algum interesse nos mesmos. Por último a exposição, também, poderá não ser intencional.

Um terceiro factor que poderá influenciar os efeitos de herança será a dimensão da audiência total disponível, ou seja, o número total de telespectadores expostos ao *media* em determinado momento. Uma vez que as mesmas pessoas estão disponíveis e veem televisão em períodos adjacentes, tem sido difícil separar as suas influências. Tiedge e Ksobiech (1986) recorreram a dados de *shares*, em vez dos valores de *ratings*, para assim tentar controlar a influência dos factores de disponibilidade nos efeitos de herança. Os *shares* são menos sensíveis relativamente a diferentes níveis de audiência total entre programas adjacentes e, conseqüentemente, será menos provável produzirem efeitos de “Lead_in” decorrentes exclusivamente da disponibilidade da audiência.

Segundo *Walker (1988)*, a tipologia do programa, também, é um factor com alguma relevância. Refere que a comédia corresponde à tipologia que maximiza esta tendência, enquanto que os programas de mistério, suspense e variedades são os menos eficazes. Esta teoria do comportamento televisivo entre programas adjacentes é válida, num contexto em que existe um número reduzido de opções, mas hoje em dia o processo de selecção poderá ser cada vez mais complexo. Verifica-se em alguns mercados um crescente número de alternativas e de programas da mesma tipologia e o público tem dificuldade em identificar todas as opções disponíveis. Por outro lado, a evolução dos dados agregados que tem fundamentado estas teorias é insensível às variações individuais. Consequentemente, coloca-se a questão de conseguir determinar com alguma exactidão o número de opções disponíveis. O telespectador individual poderá ou não ter acesso à TV Cabo, vídeo ou simplesmente poderá não estar a par de todas as opções disponíveis em determinado momento.

Cooper (1993) conseguiu ultrapassar esta dificuldade e não observou um decréscimo nos valores das correlações de valores de *ratings* de programas adjacentes com um aumento no número de opções na interface dos programas em análise. O público que tem acesso a um maior número de opções e com uma maior facilidade em seleccioná-las, transita de canal para canal entre programas ou durante os programas, anulando um efeito identificável.

O trabalho realizado por *Walker (1988)* permitiu concluir que com a penetração do controle remoto, vídeo e da TV Cabo as estratégias baseadas nos efeitos de herança terão cada vez menos sucesso, pois o público será cada vez mais activo nas suas selecções. Defende que quando um canal consegue manter os seus níveis de audiência constantes ao longo de determinado horizonte temporal, a tendência será para concluir que existe um considerável efeito de herança entre os programas, o que não será necessariamente verdade. Salienta que quando um programa conclui a maior parte da sua audiência muda de canal, mas simultaneamente origina-se um fluxo inverso vindo dos restantes canais. Em média um programa poderá herdar metade da audiência do programa anterior (*Webster, 1985*).

Os trabalhos que incidem nesta temática deveriam ser realizados com dados individuais, mas a dificuldade de acesso a esta tipologia de dados tem-no impedido. Os autores dos trabalhos identificados recorrem a indicadores acumulados de “*ratings*” que caracterizam e quantificam a duplicação das audiências entre suportes publicitários.

A análise efeitos de herança, por si só, poderá não permitir prever e explicar correctamente a evolução dos valores dos *ratings* programáticos. Será necessário, por exemplo, ter em consideração os comportamentos de lealdade ao canal, tipológica e programática. Da sua interacção resultará o valor da audiência de um programa, desconhecendo-se como os diversos comportamentos estruturais influenciam a evolução dos principais indicadores de audimetria e como se condicionam entre si. Na actualidade, também, existem algumas dúvidas a respeito da validade das teorias estruturais.

III.6.2.2. A lealdade programática

O segundo factor que influencia os *ratings* programáticos corresponde aos comportamentos de “repeat viewing”, que, também, se podem designar de lealdade aos programas. Os trabalhos desenvolvidos nesta área têm procurado caracterizar a tendência que os telespectadores apresentam no sentido de seleccionar, ao longo de um horizonte temporal, episódios consecutivos de uma série programática.

Para *Wakshlag (1985)*, seria de esperar elevados níveis de duplicação das audiências, uma vez que as preferências por determinados programas incentivam a audiência a ver os episódios subsequentes e estes últimos apresentam valores de *ratings* com uma dimensão pouco variável no decorrer do tempo (*Barwise, 1986; Ehrenberg & Wakshlag, 1987*). Porém, análises mais recentes demonstraram que os níveis de duplicação são muito mais reduzidos do que seria de esperar. Por exemplo, em Inglaterra onde o número de canais não é muito elevado, apenas metade da audiência assiste a episódios consecutivos de uma série ou telenovela (*Goodhart, Ehrenberg & Collins, 1987*).

Como causas para esta tendência, sugeriu-se a disponibilidade da audiência (*Barwise, Ehrenberg & Goodgardt, 1982; Ehrenberg & Wakshlag, 1987*) e algumas características de planeamento e de estruturação das grelhas (*Goodhart, Ehrenberg & Collins, 1987; Kameroner, 1987; Webster & Wang, 1992*). *Ehrenberg e Wakshlag (1987)* justificaram os reduzidos níveis de duplicação com a irregularidade que se verifica no consumo televisivo. No mínimo 40% dos telespectadores não estará a ver a televisão quando se transmite um episódio de um programa semanal. Portanto, quando os *ratings* apresentam alguma estabilidade no decorrer do tempo, grande parte da sua audiência não será a mesma.

Outros trabalhos (*Barwise, 1986; Barwise et al., 1982; Ehrenberg & Wakshlag, 1987; Webster & Wang, 1992*), também, tentaram relacionar a evolução dos índices de duplicação com a audiência de um programa. Verificaram que quanto mais popular for um programa maior será a tendência ser seleccionado com uma maior regularidade.

Segundo *Webster e Wang (1992)*, os índices de lealdade programática são menores no horário nobre devido aos factores de planeamento. Durante o dia é transmitido um maior número de programas com formato de séries, por exemplo, as telenovelas, assim como se verifica uma maior regularidade no padrão de consumo semanal e na audiência total disponível. Certas tipologias programáticas contribuem para uma maior duplicação das audiências, por exemplo, a estrutura narrativa das telenovelas. Entre as restantes tipologias, alguns autores, não conseguiram identificar diferenças significativas ao nível da lealdade programática (*Barwise, 1986; Barwise, Ehrenberg & Goodhart, 1982; Headen, Klompmaker & Rust, 1979; Webster & Wang, 1992*).

Na actualidade, a existência de um crescente número de alternativas e a sua própria estrutura poderá reduzir os valores dos índices de duplicação. Os reduzidos valores que se esperam na actualidade demonstram a existência de um público, talvez, cada vez mais activo, menos leal, que cada vez vê menos televisão e menos dos mesmos programas. A nível nacional, os portugueses são na realidade consumidores assíduos de televisão, por exemplo, em 2010, viram em média 3H30M de televisão (*Ferreira, Reis*

& Santos, 2011). Por conseguinte, será necessário avaliar a validade de algumas conclusões referidas em trabalhos internacionais, uma vez que estes comportamentos condicionam as abordagens de previsão a adoptar.

III.6.2.3. A lealdade ao canal

Os trabalhos desenvolvidos no âmbito da lealdade ao canal pretendem verificar se o público apresenta alguma tendência para seleccionar consecutivamente programas de determinado canal. Também, se pode designar de análises de duplicação das audiências num mesmo canal, uma vez que os trabalhos têm utilizado, essencialmente, dados agregados para avaliar se os telespectadores assistem a vários programas de um único canal.

Poucos trabalhos têm demonstrado a que nível os factores relativos aos canais influenciam a exposição. Estas variáveis podem ser influenciadas pela configuração do mercado, TV Cabo, vídeos e outras tecnologias de *media* (Cooper, 1993). Uma vez seleccionado o canal a utilização de técnicas de planeamento poderá sustentar a probabilidade de ocorrer uma certa inércia de programa para programa. Esta inércia será influenciada pelas restantes opções disponíveis e pelo seu conhecimento.

Para Cooper (1993), o conhecimento que se detém da influência do canal nas escolhas programáticas sofre das mesmas limitações encontradas na investigação a respeito dos efeitos de herança. O que se denomina de lealdade ao canal podem ser efeitos de herança disfarçados ou vice-versa. Goodhart, Ehrenberg e Collins (1975) defendem que os conceitos são diferentes, uma vez que as previsões realizadas com base nos efeitos de herança conduzem a níveis muito superiores às proporcionadas pela lealdade ao canal. Isto poderá ocorrer porque as análises realizadas têm-se concentrado apenas no grau de duplicação das audiências num mesmo canal e a inércia entre programas sustentará o nível de duplicação das audiências. Por outro lado, o público, também, poderá seleccionar em primeiro lugar um canal e assistir aos seus programas devido ao efeito da influência do planeamento e dos efeitos de herança.

A excessiva dependência dos trabalhos face a dados diários que se baseiam na memória e que tendem a maximizar o rating de programas já estabelecidos em canais mais populares, apenas contribui para a confirmação desta fraqueza. No entanto, a lealdade ao canal parece ser relevante para a explicação dos comportamentos televisivos, especialmente num contexto de grandes recursos ao nível dos *media*. Aparentemente o controlo que uma estação detém para conseguir projectar a sua imagem poderá contribuir para um maior nível de duplicação das suas audiências.

Apesar dos canais por TV Cabo reduzirem a influência dos restantes canais, a maioria dos telespectadores costuma seleccionar em maior grau os canais generalistas (*Ferguson, 1992; Heeter, 1988*). Poucos telespectadores tiram partido da grande variedade de canais de que dispõem (*Ferguson, 1992; Ferguson & Perse, 1993; Heeter, 1985; Heeter & Greenberg, 1988; Lochte & Warren, 1989*). No entanto, diferenciam-se quanto aos canais generalistas seleccionados com uma maior frequência. Apesar da subscrição da TV Cabo aumentar o número de canais seleccionados, esta expansão é muito menor do que intuitivamente se poderia pensar a prior. *Heeter e Greenberg (1988)* defendem que a introdução da TV Cabo poderá acentuar a lealdade aos canais devido a três factores:

- Os canais por cabo especializam-se em determinados conteúdos, logo a lealdade ao canal alinhar-se-á com a lealdade a tipologias de programas;
- Existindo um custo associado aos canais por cabo é provável que os telespectadores assistam em maior grau a estes canais do que se estes estivessem isentos de custos;
- Quando o número de canais disponíveis é grande, a identificação do público com um subgrupo de canais simplifica o processo de selecção.

III.6.2.4. As exposições repetidas

A exposição repetida representa a tendência do público para assistir a programas que são retransmitidos. Os trabalhos realizados a este nível são os mais escassos na bibliografia da especialidade.

Cooper (1993) verificou que é mais provável o público assistir à primeira emissão do que às emissões subsequentes, mas os resultados variavam em função dos canais considerados. O público de programas com maiores níveis de exposição apresenta uma maior tendência para assistir a retransmissões, mas a sua proporção é inferior a um terço (*Litman & Kohl, 1992*).

III.6.2.5. A avaliação da interacção entre as diversas vertentes de análise da duplicação das audiências

Apenas foi possível identificar um trabalho que pretende definir as interdependências entre os vários comportamentos estruturais de consumo televisivo (*Cooper, 1996*). O autor apresenta uma proposta teórica, mas não validada empiricamente. Salienta que a lealdade ao canal e os efeitos de herança funcionam como “produtores” de audiências influenciando a exposição entre e através dos programas de determinado canal. Os níveis de lealdade programática e de exposições repetidas dependem das audiências dos programas anteriores e das tipologias dos programas em análise.

III.6.2.6. Os modelos de exposição por previsão da duplicação das audiências

Apenas foi possível identificar uma referência a um trabalho que procura prever as audiências totais de uma série programática através da previsão do número de telespectadores duplicados e não duplicados entre os diversos suportes que não devem transcender as 6 a 12 emissões e que se encontra descrito em *Wyndham (1994)*. As análises apresentam margens de erro que podem alcançar os 30%, apesar do menor horizonte temporal que uma série com este número de episódios abrange e considerando como variável de previsão a atractividade que os telespectadores manifestam a respeito dos seus episódios iniciais.

III.6.3. As abordagens alternativas de construção dos modelos de exposição

Os planos de *media*, também, podem ser construídos segundo perspectivas que transcendem a utilização dos principais indicadores agregados e acumulados de audimetria. Em *Zarate (1997)*, consta uma crítica aos modelos que recorrem apenas aos inquéritos e salienta a necessidade da previsão da exposição ser alcançada através da interpretação e caracterização da estrutura dos comportamentos de consumo televisivo, sendo as bases de audimetria essenciais neste processo. A autora refere que, do ponto de vista metodológico, existe a necessidade de recorrer à classificação de dados e propõe uma sequência de análises que inclui a aplicação da análise Discriminante Linear, análise de *Clusters* e da análise de Correspondências, as quais devem ser complementadas com abordagens qualitativas de caracterização do consumo televisivo. As teorias explicativas da exposição televisiva contribuem para a identificação das variáveis que determinam o consumo televisivo.

A sequência de análises proposta contribui, principalmente, para uma selecção dos suportes que será função do grau de exposição ao painel e não se pretende a previsão de um valor concreto, ainda que de foram aproximada.

As metodologias de Estatística Multivariada na realidade não correspondem às mais adequadas para a análise de bases de dados de audimetria, ou de quaisquer outras que, frequentemente, surgem no meio industrial, uma vez que apresentam um elevado número de registos, e por vezes anomalias, duplicações, inconsistências ou omissões, para além da não lineariedade das relações entre as variáveis. A Estatística clássica não permite e não é aplicável à análise de inúmeras bases de dados sobre as quais devem incidir as múltiplas áreas metodológicas enquadradas na área do *Data Mining*.

Segundo *Branco (2010)*, a diferença mais marcante entre a estatística e o *Data Mining* é a grande dimensão e complexidade dos dados analisados pelo DM, contrariamente ao que sucede em Estatística que no geral apenas se ocupa de pequenos conjuntos de dados estáticos, com uma pequena quantidade de anomalias, constituindo uma amostra com vista a responder a questões concretas relativas à população de onde a amostra foi retirada. O autor salienta que o uso de modelos assente em hipóteses não se adequa aos dados analisados em DM e que a estatística deve incidir na análise de dados que foram concebidos principalmente como dados experimentais resultantes de algum delineamento, ou pelo menos obedecendo a algum processo de recolha previamente estipulado. O DM permite a identificação de padrões, tendências inesperadas ou interacções existentes em grandes bases de dados, promovendo uma maior percepção do processo subjacente

Em *Zarate (1997)*, podemos encontrar algumas referências a modelos de previsão das audiências de programas, mas que na sua maioria recorrem à análise estatística multivariada de alguns indicadores de audimetria ou às metodologias mais tradicionais de análise de séries temporais e às abordagens de previsão qualitativas. As análises de previsão qualitativas, também, são insuficientes e pouco adequadas ao planeamento de *media*. As abordagens qualitativas apresentam inúmeras desvantagens, sendo as perspetivas mais utilizadas: as sondagens de intenção de ver e os painéis de opinião.

As previsões por juízos de valor acarretam um sobreajustamento aos dados, uma vez que existe uma tendência para serem identificados padrões sistemáticos imaginários onde apenas existe ruído e tendem a extrapolar esses padrões. Os painéis apresentam uma limitada capacidade de processamento dos dados e simplificam demasiado as questões. Esta metodologia é mais adequada em contextos que envolvem um reduzido número de variáveis e de dados. Podem ter em atenção factores qualitativos, mas que por sua vez não são interpretados de forma objectiva e neutra. Os resultados não são generalizáveis uma vez que dependem do “Know-How” de um utilizador e de juízos de valor que são subjectivos.

Inclusive numa perspectiva de previsão da repercussão de observações atípicas, como, por exemplo, quando são transmitidos os eventos desportivos, a sua utilização não conduz a bons resultados (*Nikolopoulos, 2010*).

Do ponto de vista das sondagens de intenção de ver os resultados não são fiáveis pela falta de objectividade do que é declarado pelos telespectadores. As selecções dependem de inúmeros factores que os telespectadores desconhecem ou, não, consideram em determinado momento. Os custos e o tempo necessário para a análise dos dados, também, representam factores que penalizam esta abordagem.

CAPÍTULO IV

METODOLOGIA

Este capítulo aborda a descrição do tipo e método de investigação adoptado e da estrutura metodológica das análises propostas, referindo-se os resultados esperados e as respectivas condições de aplicabilidade. Privilegiam-se algumas abordagens enquadradas na Análise Simbólica e no *Data Mining* Proposicional, Temporal e Sequencial, justificando-se a necessidade da sua utilização. Caracteriza-se sinteticamente a fonte de dados de audimetria, as diversas bases de dados que foram construídas e a amostra.

IV.1. INTRODUÇÃO

Um trabalho de investigação visa a identificação de respostas para determinados problemas, mediante a utilização de procedimentos científicos racionais e sistemáticos, que assegurem a exactidão da informação (Gil, 1999). De acordo com Caldeira (2000), a pesquisa pode ser desenvolvida num enfoque interpretativo, ou segundo as características do Positivismo, Pós-Positivismo, Realismo Crítico e Construtivismo. Para Easterby-Smith et al. (1999), destacam-se duas principais tradições em ciência: o Positivismo e a abordagem Fenomenológica, Qualitativa ou Interpretativa, às quais se associam determinadas suposições e implicações metodológicas. O positivismo de Auguste Comte defende a observação dos fenómenos, o desenvolvimento da experiência, a descoberta de leis, quer dizer relações que possam permanecer constantes entre os fenómenos observáveis no mundo natural e social, opondo-se ao racionalismo e ao idealismo. Este nega à ciência a possibilidade de investigação das causas dos fenómenos naturais e sociais, como consequência da sua inacessibilidade e falta de utilidade.

O Positivismo defende que o conhecimento científico é o único, que permite alcançar um conhecimento verdadeiro. Por outro lado, na Fenomenologia a tarefa de um cientista social não é levantar factos e medir frequências de determinados padrões de comportamento, mas apreciar os significados que as pessoas atribuem à sua experiência. Defende que a realidade não é objectiva e exterior ao homem, mas socialmente construída.

Uma investigação deve-se aproximar de determinada corrente de pensamento científico, e este trabalho apresenta características, que em termos metodológicos, se aproxima do Positivismo. A realidade nesta perspectiva é estável, observável, objectiva e mensurável e pressupõe que os fenómenos estão sujeitos a um conjunto de leis invariáveis, tanto no mundo humano, como no natural o que suscita grandes dúvidas para inúmeros cientistas sociais. Também rejeita que as qualidades humanas transcendem o alcance da compreensão científica. Na abordagem Positivista procuram-se regularidades e relações entre os elementos, que constituem determinado sistema e que possam apresentar alguma estabilidade.

De acordo com *Godoy (1995)*, os métodos quantitativos são os mais indicados nas investigações com uma orientação filosófica Positivista, onde se defende que a objectividade sómente é possível quando o trabalho é realizado por recurso a instrumentos padronizados, supostamente neutros, assegurando assim que sejam realizadas generalizações com precisão e objectividade. Os métodos qualitativos são considerados como sendo mais indicados para as investigações de perspectiva interpretativa ou crítica.

Seguindo esta perspectiva, na presente tese, adopta-se uma metodologia quantitativa, uma vez que se pretende caracterizar a exposição televisiva e a evolução dos indicadores que dela resultam, com o objectivo de identificar regularidades nos dados que permitam contribuir para a futura previsão das audiências televisivas.

Pretendemos identificar quais são as melhores perspectivas metodológicas para a construção dos modelos de exposição, após uma maior percepção dos factores que condicionam a evolução dos “ratings” e após a avaliação da variabilidade que apresenta a sua evolução.

As análises descritivas permitem compreender melhor o contexto da investigação. O principal objectivo das análises consiste na descrição e sempre que possível na explicação da evolução dos dados e na procura de regularidades, que possam ser generalizadas e viabilizem a previsão dos dados. Na metodologia quantitativa a realidade é assumida como objectiva e, logo, adopta-se uma visão Positivista da ciência.

A classificação das Ciências em Exactas e Inexactas divide-as de acordo com o grau de precisão das descrições e previsões realizadas, com base nos modelos e nas teorias científicas.

As Ciências Exactas, como, por exemplo a Física, geralmente conseguem proporcionar resultados que apresentam um elevado grau de precisão a respeito dos sistemas em análise, enquanto nas Ciências Sociais, sendo inexatas as previsões são geralmente apenas direccionais e, logo, não exactas.



As Ciências Humanas apresentam um carácter muito menos preciso e por isso, impera a perspetiva explicativa dos comportamentos e dos dados observados e eventualmente previsões meramente direccionais e probabilísticas.

Fig. IV.1 – O Método indutivo e o método dedutivo.
Fonte: Elaborado pela autora.

A investigação científica depende de um “conjunto de procedimentos intelectuais e técnicos” (Gil, 1999, p.26) para que os seus objectivos sejam alcançados: os métodos científicos. O método científico é o conjunto de processos que se devem utilizar na investigação. Os métodos que fornecem as bases lógicas à investigação são os seguintes: o dedutivo, indutivo, hipotético-dedutivo, dialético e fenomenológico (Gil, 1999). O método indutivo considera que o conhecimento é fundamentado na experiência, não levando em conta princípios pré-estabelecidos.

No raciocínio indutivo a generalização deriva de observações de casos da realidade concreta. As constatações particulares levam à elaboração de generalizações (Gil, 1999). Na presente tese, o quadro metodológico ou desenho da investigação tem como referência o método indutivo, sendo essencial recorrer a grandes bases de dados, que permitam caracterizar a relação dos telespectadores com o televisor. As metodologias de DM permitem a identificação de eventuais regularidades nos dados históricos, com base na análise de uma amostra que deverá ser o mais representativo do universo em estudo e que posteriormente são generalizadas para o universo televisivo e para um período futuro.

IV.2. MODELO DE INVESTIGAÇÃO

A pesquisa é um conjunto de acções, propostas para encontrar a solução para um problema. O modelo de investigação depende do estágio de desenvolvimento em que se encontra o conhecimento sobre determinada temática. Os trabalhos podem ser classificados em exploratórios, descritivos e causais. De acordo com *Espinosa (2006)*, os trabalhos exploratórios procuram indícios sobre uma determinada questão, analisando as possíveis alternativas de decisão e identificando as variáveis mais relevantes para o problema em análise. Visam proporcionar uma maior familiaridade com o problema, tornando-o mais explícito e permitindo a construção de hipóteses. Os estudos descritivos descrevem uma temática e tratam da sua decomposição. Por último, os estudos causais procuram obter evidências sobre uma relação de causa-efeito entre as variáveis.

Segundo *Gil (1999)*, a pesquisa exploratória tem como objectivo desenvolver, esclarecer e modificar conceitos e ideias, com base na formulação do problema ou hipóteses para estudos posteriores. Assume alguma flexibilidade, versatilidade e podem incidir em dados secundários. A pesquisa explicativa caracteriza-se pela identificação das causas que contribuem para a ocorrência do fenómeno e aprofunda o conhecimento da realidade, porque explica o porquê dos factos. A pesquisa descritiva pretende descrever as características de determinada população, fenómenos ou o estabelecimento de relações entre as variáveis.

As análises propostas enquadram-se maioritariamente no âmbito da investigação exploratória e descritiva. Correspondem a análises exploratórias, uma vez que se pretende a identificação das perspectivas de análise que podem conduzir aos melhores resultados quando se pretende a construção dos modelos de exposição. Tenta-se identificar as variáveis mais relevantes e avaliam-se as potencialidades de diferentes perspectivas metodológicas que devem ser função dos contextos cujos devem ser previamente identificados nos dados. Atendendo às dificuldades que estes modelos colocam ao sector televisivo, pretende-se proporcionar um maior conhecimento sobre a evolução dos dados de audimetria, contribuindo para a posterior construção dos modelos de exposição.

As análises exploratórias visam identificar cursos relevantes de acção ou obter dados adicionais, antes que se possa desenvolver uma abordagem. Devem permitir uma maior compreensão do problema que determinado investigador enfrenta, e que neste contexto se traduziu em compreender a evolução dos indicadores agregados e acumulados de audimetria. A construção de um modelo estrutural explicativo da exposição televisiva e a análise dos “gross measures” fundamenta as propostas descritas no âmbito da construção dos modelos de *media*. As análises realizadas permitiram obter critérios, para desenvolver uma abordagem do problema e estabelecer prioridades para pesquisas posteriores. As propostas descritas diferenciam-se em função do contexto em que são aplicáveis e foram hierarquizadas em função da sua complexidade e dos recursos que envolvem.

Foram identificadas as variáveis e as relações-chave que devem ser consideradas em análises posteriores.

Na nossa investigação, constam análises descritivas que caracterizam a exposição televisiva sob várias vertentes, recorrendo sucessivamente a diferentes tipologias de dados, que permitem ir aprofundando as análises e explicando a evolução e variabilidade dos principais indicadores de audimetria. Analisa-se a repercussão das variáveis estruturais de audimetria na evolução dos “*ratings*”, identificando as mais relevantes e as respectivas interdependências. As perspectivas metodológicas que viabilizam o investimento em publicidade televisiva dependem do prévio conhecimento e compreensão da evolução dos dados.

A pesquisa descritiva pretende observar sistemas, descrevendo-os, classificando-os e interpretando-os. Também, se pode interessar pelas relações entre as variáveis, mas não tem o compromisso de os explicar, embora sirva de base para tal explicação. Normalmente baseia-se em amostras grandes e representativas, correspondendo neste caso às bases de audimetria que apresentam uma elevada dimensionalidade e caracterizam o comportamento de um painel considerado representativo do conjunto de lares portugueses que têm televisor.

Os estudos mais utilizados neste tipo de pesquisa são o longitudinal e o transversal. Um painel de audimetria viabiliza uma análise longitudinal da exposição televisiva, e por conseguinte constitui o principal instrumento de previsão e de planificação que existe na actualidade. As análises foram realizadas com base em uma amostra, que abrange equivalentes períodos no decorrer de vários anos, permitindo, assim, a análise da evolução dos dados, que determinam o investimento publicitário. Nos seguintes capítulos, descreve-se esta fonte de dados e as características da amostra seleccionada.

IV.3. DESCRIÇÃO DA TECNOLOGIA DE AUDIMETRIA

Os principais instrumentos de medição das audiências televisivas são os estudos de memória, os diários de audiências, os estudos coincidentais e a audimetria.

Os sistemas de informação podem ser classificados em audimetria activa, audimetria passiva, audimetria portátil e audimetria interativa. Os dados de audimetria disponibilizados para a realização das análises resultam da tecnologia de audimetria activa, onde um sistema centrado em painel de 750 lares regista a exposição individual. Cada indivíduo que integra o painel deverá assinalar a sua presença e as suas selecções através de um número que lhe é atribuído no telecomando, sendo esta informação enviada via telefone para a Marktest Audimetria. Nos dados dos ficheiros base de audimetria estão representados os instantes em que cada telespectador liga, desliga ou muda de canal, sendo posteriormente validados, formatados e ponderados (*Fig. IV.3.1*).



Fig. IV.3.1 – Representação do processo de audimetria de medição das audiências televisivas.
Fonte: Elaborado pela autora.

A partir dos dados dos ficheiros base de audimetria, também designados de “*raw data*”, são construídas duas bases de dados de índices de audiências que caracterizam diversas vertentes da exposição televisiva aos programas, espaços comerciais e inserções, ou seja, anúncios publicitários com alguns segundos de duração. A base de dados *Telereport* permite a análise das audiências de programas e de blocos de anúncios, estando disponíveis os designados “*gross measures*” e os indicadores acumulados que permitem a análise da audiência partilhada por diversos suportes. A base de dados *Publitem* integra informação referente a cada uma das inserções publicitárias transmitidas no decorrer de um espaço comercial.

O método de amostragem probabilístico adoptado para a constituição do painel de audimetria é o estratificado desproporcional. O painel é caracterizado em função das seguintes variáveis sociodemográficas: região geográfica, mas apenas a nível de Portugal Continental, sexo, faixa etária, classe social, situação no lar, quer dizer, se é dona de casa e responsável pelas compras do lar e por último se o lar tem TV Cabo. No *Apêndice A* consta uma descrição mais exaustiva dos diversos métodos de medição das audiências televisivas.

IV.4. DESCRIÇÃO DAS BASES DE DADOS

Os anunciantes ou as agências de meios que os representam e planificam as campanhas publicitárias, no geral apenas têm acesso às bases de dados de índices de audiências (*Telereport* e *Publitem*), sendo esta a informação disponibilizada para a realização das análises propostas.

A partir dos *softwares Telereport* e *Publitem* foi necessário construir uma diversidade de bases de dados (*Fig. IV.4.1*), que se caracterizam detalhadamente em cada um dos capítulos de apresentação de resultados, que constam no *Apêndice F*, nomeadamente uma base de dados para a análise da evolução dos principais indicadores programáticos e uma segunda referente aos dados comerciais.

A análise dos indicadores acumulados implicou a construção de diversas bases de dados, entre as quais, um base de indicadores de lealdade programática, outra referente à análise dos efeitos de herança e de fluxos entre canais e uma terceira onde constam diversos índices sociodemográficos. A estrutura inicial dos dados inviabiliza a realização das análises propostas.

Após a extração e agregação dos dados foi necessário proceder à sua filtragem, correcção e ao cálculo de inúmeras variáveis, que não constam das bases de dados originais.

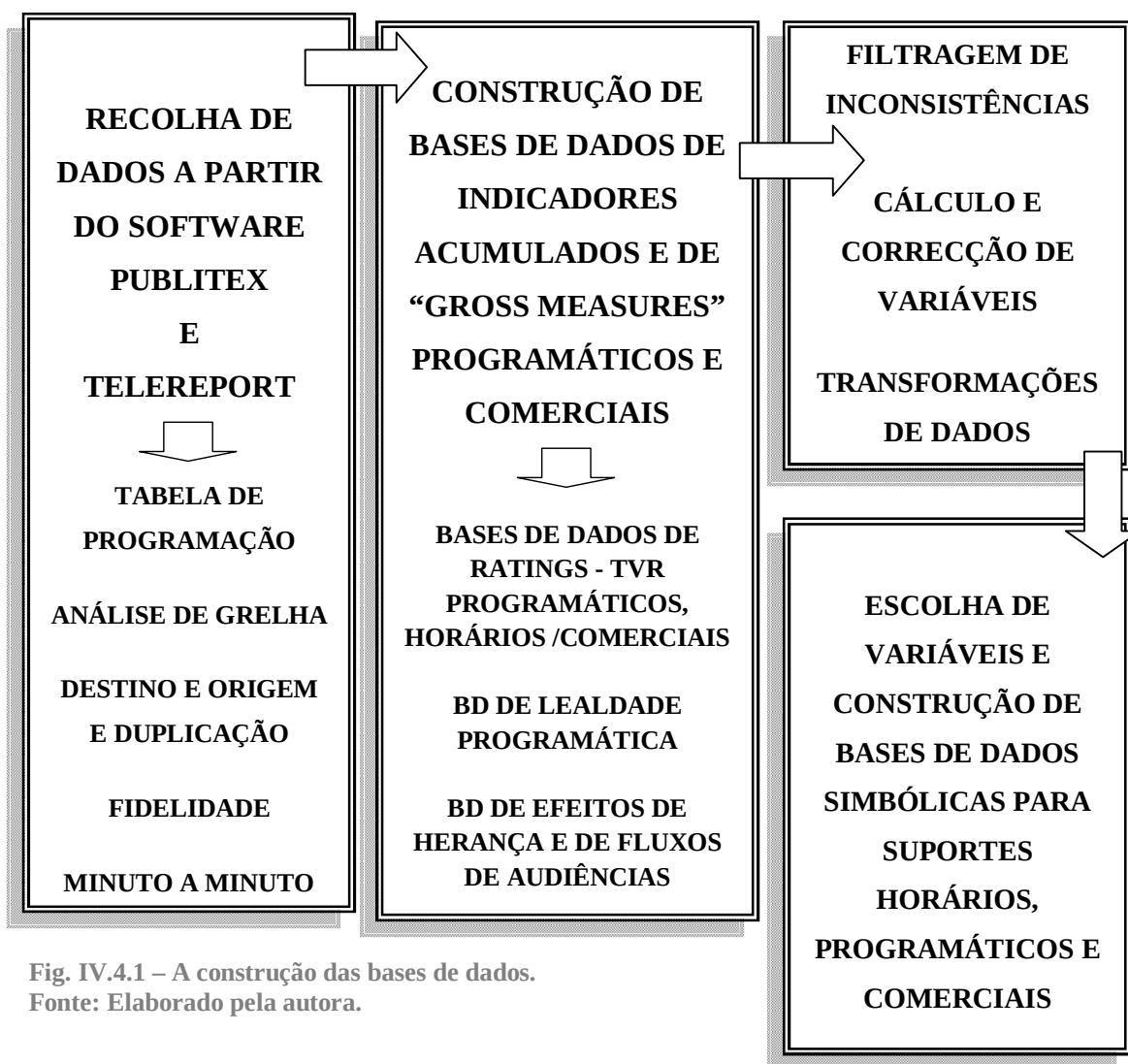


Fig. IV.4.1 – A construção das bases de dados.
Fonte: Elaborado pela autora.

Foram consideradas as bases *Telereport* de tabelas de programação, análise de grelha, destino e origem, fidelidade, minuto a minuto e de duplicação de audiências. Concluiu-se com a construção das bases de dados simbólicas, recorrendo ao *software* SODAS versão 2 (*Diday & Noirhomme-Fraiture, 2008*) e ao Microsoft Office ACCESS 2003 e 2010.

IV.5. HORIZONTE TEMPORAL

Um estudo pode ser seccional quando os dados se referem apenas a um determinado momento e variáveis, e pode ser longitudinal quando os dados se referem a vários momentos no decorrer do tempo. Nas análises foram utilizados dados de audimetria referentes a diversos períodos, logo, é um estudo longitudinal. Sendo o objectivo das análises a construção dos modelos de exposição necessários ao planeamento de *media* será inevitável recorrer à audimetria, uma vez que permite ao contrário dos restantes instrumentos, uma medição rápida, contínua, detalhada e mais objetiva da exposição televisiva.

Os instrumentos de medição das audiências que apenas viabilizam uma recolha pontual de dados e que incidem num reduzido horizonte temporal, como, por exemplo os inquéritos e os diários de audiências, não constituem instrumentos de previsão. Apenas permitem conhecer a audiência num período que corresponde ao do estudo, proporcionando uma visão estática do consumo televisivo, e por conseguinte inadequada quando os comportamentos apresentam uma considerável variabilidade do decorrer do tempo e não podem ser generalizados entre sucessivos períodos de curta duração. Os dados que resultam de inquéritos não são objectivos, impedem a quantificação da exposição histórica e futura, assim como o cálculo de informação que é essencial para o investimento em publicidade televisiva. Consequentemente, os restantes métodos de medição das audiências apenas se utilizam, actualmente, como um complemento. As análises realizadas incidem num horizonte temporal abrangente e recente, onde se verifica uma significativa alteração neste sector com a crescente

consolidação da posição da TVI no panorama audiovisual nacional. Consideraram-se neste trabalho, os anos compreendidos entre 2000 a 2010 e para cada ano considerado as semanas que decorrem desde 15 de Março a 15 de Maio, uma vez que representa um período considerado característico a nível médio do consumo televisivo. Em alguns contextos apenas se apresentam os resultados para o período que apresenta uma maior instabilidade no desempenho dos canais e que permite validar a adequabilidade e precisão de algumas das metodologias adoptadas, ou seja, desde 2000 a 2004 (*ver Apêndice B*).

IV.6. UNIDADE DE ANÁLISE

A unidade de análise corresponde a programas, intervalos horários, blocos de anúncios ou inserções comerciais. Nas análises simbólicas, as observações representam um conjunto de unidades das referidas características. Considera-se a totalidade da emissão diária, e não apenas as faixas horárias da noite, onde incide a maioria dos trabalhos que constam da bibliografia internacional.

Privilegiam-se os dados dos canais generalistas *RTP1*, *Canal 2*, *TVI* e *SIC* pelo seu maior valor “comercial”. As campanhas publicitárias incidem em maior grau nos canais aos quais tem acesso a totalidade do painel. Contudo, considera-se a repercussão do desempenho dos canais não generalistas, ou seja, por TV Cabo e satélite, assim como do vídeo. A descrição da programação, no *software Telereport*, também apenas se encontra disponível para os canais generalistas. Nas análises sociodemográficas considera-se a totalidade das variáveis que constam das bases de dados de audimetria.

V.7. CARACTERIZAÇÃO E JUSTIFICAÇÃO DAS ÁREAS METODOLÓGICAS ADOPTADAS

A evolução verificada ao nível dos instrumentos metodológicos não tem acompanhado a crescente complexidade, que caracteriza o consumo televisivo.

Apenas o desenvolvimento de novos métodos quantitativos permitirá a análise da exposição televisiva no futuro (Weber, 2002). Assim sendo, face aos objectivos propostos e de acordo com as lacunas identificadas no estado da arte dos modelos de *media* de exposição, as metodologias a serem utilizadas devem proceder das áreas mais adequadas às características destas bases de dados. Privilegiou-se a Análise Simbólica (Bock & Diday, 2000) e o *Data Mining* (Han & Kamber, 2001), complementando-se com a Estatística Multivariada (Reis, 1997).

A previsão apenas será viável, através da identificação de regularidades nos dados e recorrendo a análises de indução. As metodologias de *Data Mining* são, conseqüentemente, as mais adequadas para a identificação de regularidades e interacções não lineares, em bases de dados que apresentam uma elevada dimensionalidade. As metodologias, técnicas e algoritmos visam não só a área da Previsão e Regressão linear ou não Regressão linear, mas, também, a segmentação de dados, ou seja a classificação, a análise de agrupamentos e a redução da dimensionalidade recorrendo a Métodos Factoriais e a análises *Multidimensional Scaling* (MDS).

A Análise Simbólica representa a extensão das metodologias multivariadas de análise dados e de *Data Mining*, para a análise de bases de dados onde uma observação não consiste num único valor nominal ou numérico. Conseqüentemente, permite uma representação dos dados de audimetria em estruturas mais ricas face às tradicionais, viabilizando a realização de análises mais adequadas, à caracterização da evolução temporal dos índices de audiência e ao desenvolvimento das correspondentes análises explicativas, que eventualmente também podem apresentar capacidade de previsão.

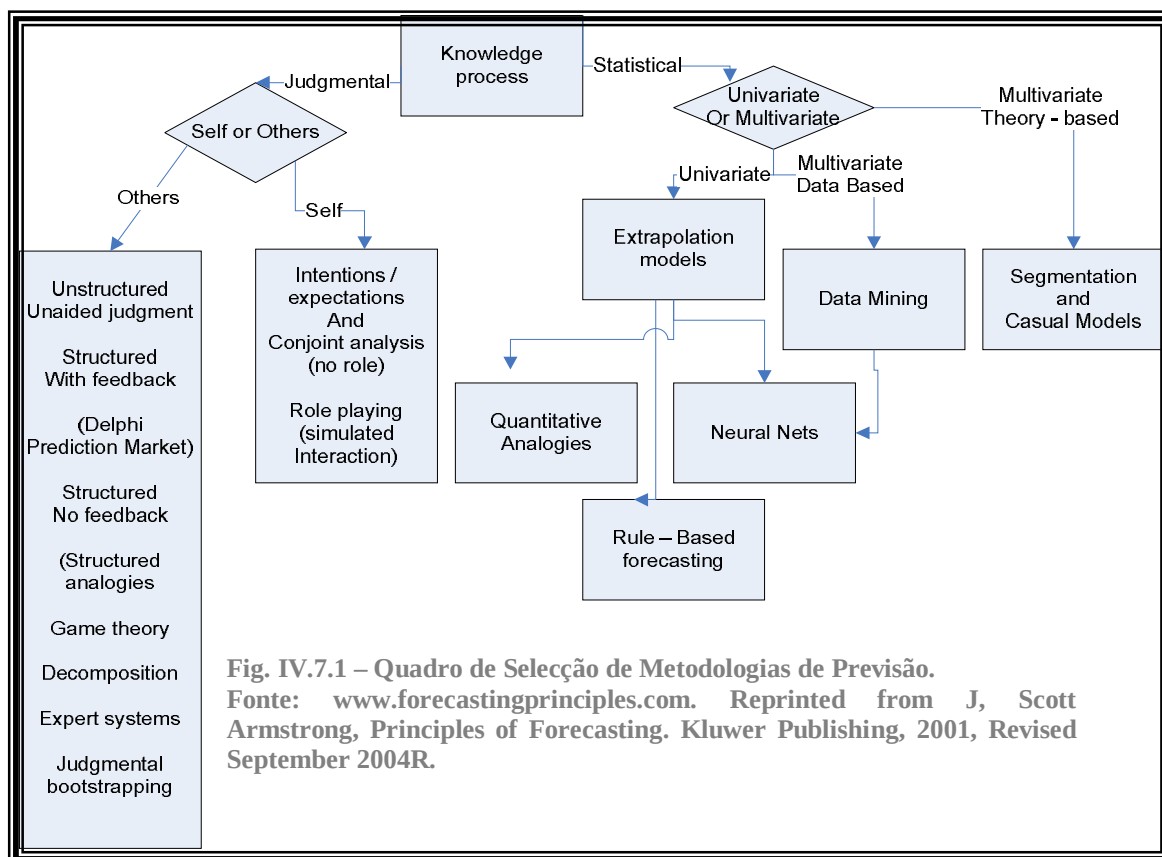
As análises multivariadas poderiam ser realizadas recorrendo a variáveis que traduzem médias, medianas ou variâncias das variáveis iniciais, mas esta opção poderá conduzir a resultados menos claros face à utilização das variáveis simbólicas, que podem ser classificadas nas seguintes tipologias: intervalo, histograma, diagrama, multivalor, taxonómica e hierárquicas.

A Análise Simbólica permite a agregação dos dados em unidades estatísticas mais relevantes, mas cada análise requer uma atenção especial quanto à definição da tipologia de variável mais adequada.

As metodologias associadas às inúmeras áreas de *Data Mining* visam a identificação de relações, padrões e tendências significativas e inesperadas em grandes bases de dados, captando a não-linearidade das relações. O *Data Mining* encontra-se em forte expansão, recebendo contributos de inúmeras áreas, como, por exemplo, do *Machine Learning*, Estatística, Optimização, Matemática Computacional, Teoria de Algoritmos, Visualização e Bases de Dados. As áreas que *a priori* parecem mais relevantes para a análise de dados de audimetria, correspondem ao *Data Mining* Temporal (Laxman & Sastry, 2006; Wiu, Agrawal & Abaddi, 2000), *Data Mining* Sequencial (Xin, Han, Cheng & Yan, 2007), *Data Mining* Multirrelacional (Dzeroski & Lavra, 2001; Dzeroski, 2005) e as metodologias proposicionais de Regressão não linear ou os algoritmos bayesianos (Kennett, Korb & Nicholson, 2001) e as Regras de associação (Rauch & Simunek, 2005; Zaki, 2001).

A Figura IV.7.1 e a Figura IV.7.2 ilustram as metodologias de previsão e explicativas mais adequadas em diferentes contextos. Em primeiro lugar subdividem-se em métodos quantitativos ou baseados em juízos de valor. Ambos são internacionalmente utilizados para a previsão das audiências televisivas, produzindo melhores resultados os métodos quantitativos, que são essenciais para o planeamento de *media*. Os métodos não quantitativos, como, por exemplo, as sondagens de intenção de ver e os painéis de opinião têm sido utilizados em contextos de maior variabilidade dos índices de audiências, mas, não, proporcionam resultados satisfatórios e informação, que é essencial para a actividade de planeamento de *media*. As sondagens de intenção de ver não são objectivas, uma vez que apenas nos proporcionam a informação que é divulgada pelo próprio sujeito e que corresponde ao objecto de medida. Não existe uma ausência de estímulos, que na realidade estão presentes através da influência do entrevistador e dos estímulos internos do próprio entrevistado.

Nos métodos quantitativos quando não existe conhecimento a respeito das relações existentes e estamos perante séries temporais, os modelos mais adequados correspondem aos baseados na extrapolação, assim como quando existe conhecimento e uma maior estabilidade na evolução dos dados. As restantes áreas de previsão quantitativas incluem a identificação dos cenários mais semelhantes à observação a prever, metodologias de inteligência artificial: os sistemas expertos, que podem ser desenvolvidos por recurso a Redes Neurais, entre outras abordagens, a extracção de regularidades (“rule-based forecasting” com a utilização de Regras ou Árvores de decisão de *Data Mining*), as simulações e os modelos econométricos, como, por exemplo, os modelos de Regressão linear. A *Figura IV.7.1* classifica em categorias as diversas metodologias de previsão e descreve como estas se inter-relacionam, representando as linhas a tracejado as possíveis relações existentes. As análises estatísticas são inicialmente classificadas em multivariadas ou univariadas. As análises multivariadas podem incidir na identificação de regularidades existentes nos dados ou no ajustamento de determinados modelos matemáticos, como os causais lineares ou recorrendo à classificação de dados.



As metodologias de previsão também incluem as análises de extrapolação e as metodologias de *Data Mining*, como, por exemplo as Redes Neurais ou as Regras de Decisão. Os dados de audimetria inserem-se na classificação de estatística – multivariado - univariado, sendo as metodologias de Regressão não linear de DM as mais indicadas, quando os modelos se baseiam em padrões que se pretendem identificar nos dados ou em modelos causais lineares quando existe um modelo teórico subjacente. O DM permite identificar relações entre as variáveis que podem ser, posteriormente, consideradas em modelos causais.

Os métodos quantitativos de previsão partem do pressuposto de existência de alguma estabilidade e procuram projectar no futuro os padrões que se identificam nos dados. Quando não existe a perspectiva de ocorrerem significativas alterações no contexto histórico, podemos recorrer aos métodos quantitativos, que podem ser diferenciados em causais ou não causais (Tavares *et al.*, 1996). As análises quantitativas devem, por isso, incidir em maior grau no curto e médio prazo e tem que ser avaliada a regularidade dos dados.

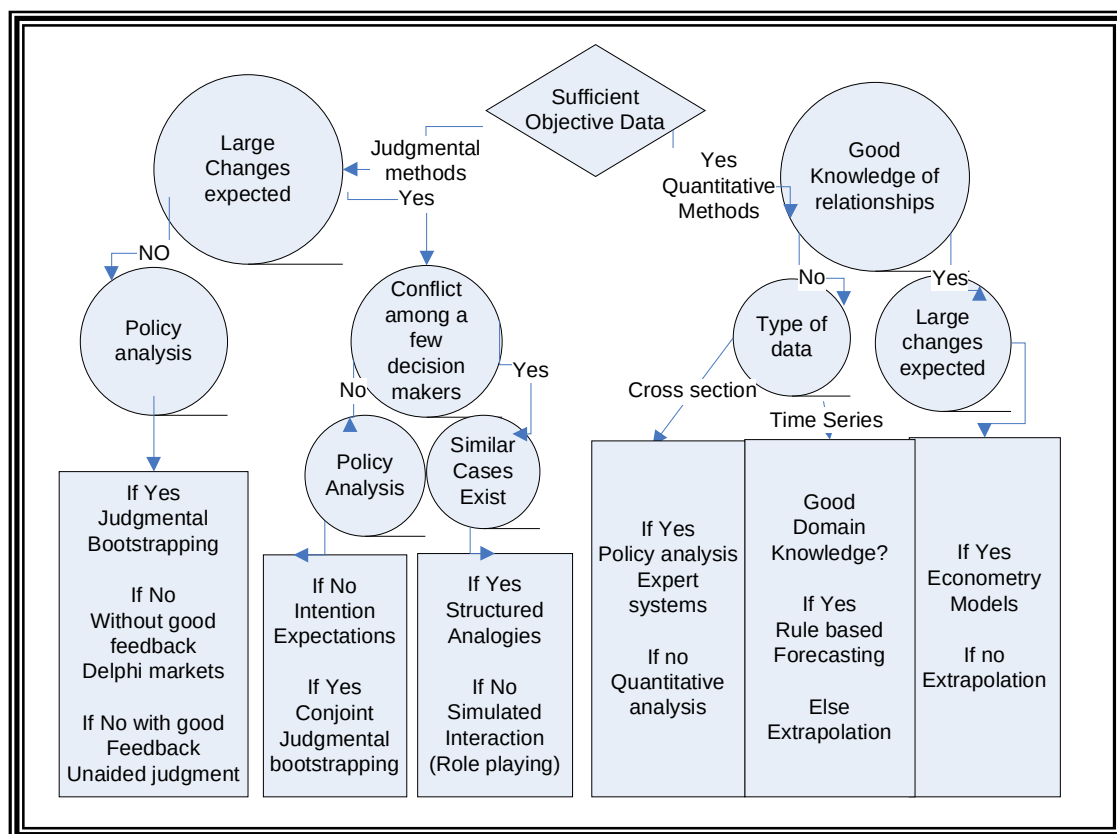


Fig. IV.7.2 – Árvore metodológica para a selecção de metodologias de previsão.

Fonte: www.forecastingprinciples.com.

Os métodos causais procuram relacionar a variável a prever, com as variáveis que permitem explicar a sua evolução. Neste trabalho foram realizadas análises exploratórias, que procuram avaliar as interações entre as variáveis. Os métodos não causais assentam apenas na análise de séries temporais de valores passados da variável, procurando caracterizar a sua evolução e projetando no futuro esses padrões de comportamento. Foram privilegiadas as análises multivariadas, mas também constam algumas perspectivas de análise univariada dos dados. Os métodos de *Box e Jenkins* e os Bayesianos permitem combinar a informação contida na série cronológica da variável a prever, com os dados relativos a outras variáveis explicativas, promovendo uma simbiose com os métodos causais.

Os métodos não causais estão mais vocacionados para fazer previsões no curto prazo, enquanto nos métodos causais o horizonte pode-se estender ao médio prazo. Nas previsões de longo prazo, porque a hipótese de estabilidade é mais questionável, alguns autores salientam (*Tavares et al., 1996*) a necessidade de recorrer a métodos não quantitativos, mas que no contexto da audimetria não têm conduzido a resultados satisfatórios.

Inclusive considerando um maior horizonte temporal as metodologias quantitativas devem pelo menos complementar as análises. Neste estudo pretende-se principalmente recorrer a análises multivariadas de Regressão linear ou ao equivalente algoritmo simbólico e à Regressão não linear, uma vez que permitem considerar variáveis explicativas da evolução dos dados. Contudo, uma vez que uma das questões em análise corresponde à avaliação da viabilidade, de previsão dos valores de *ratings* e das taxas médias de audiências para o total da TV, também se realizaram algumas análises de previsão de séries temporais univariadas, que apenas consideram a evolução temporal dos dados. No decorrer dos períodos horários mais regulares do ponto de vista da evolução dos *ratings*, a análise das séries temporais poderá ser relevante para a sua previsão, porém não devem ser apenas restringidas a algoritmos de previsão, mas também incluir por exemplo, a segmentação e o clustering das bases de dados.

No *Quadro IV.7.1* constam algumas metodologias de Regressão para uma ou várias variáveis dependentes. Quando se pretende a realização de análises de Regressão não Linear, geralmente recorre-se em primeiro lugar à Regressão linear com termos não lineares, ou à transformação da variável dependente, para ser viável o ajustamento de um modelo linear. Quando nenhuma destas opções conduz a resultados satisfatórios, devemos em substituição optar por metodologias de Regressão não Linear. As análises não lineares foram privilegiadas neste trabalho, atendendo à dimensionalidade das bases de dados, mas comparando a sua capacidade explicativa e/ou predictiva com o desempenho dos modelos de Regressão linear simples e múltipla, incluindo os equivalentes algoritmos simbólicos.

Salientamos algumas áreas metodológicas de Regressão não linear que podem ser identificadas na bibliografia (*Han & Kamber, 2001*), atribuindo uma maior ênfase às áreas do *Data Mining*;

- (1) Os algoritmos matemáticos que descrevem soluções como uma operação directa sobre uma nova observação, inserindo-se nesta área as Redes Neurais (*Chow & Cho, 2007; Tang, 2007*) e as soluções estatísticas não lineares;
- (2) As metodologias de distância que calculam uma solução para uma nova observação, em função do seu grau de semelhança face aos restantes dados existentes na sua vizinhança, sendo um dos algoritmos mais utilizados o KNN (*Jain, 2010*);
- (3) Os algoritmos lógicos que proporcionam soluções lógicas que consistem em expressões que são avaliadas como verdadeiras ou falsas, aplicando operadores comparativos ou booleanos às variáveis. Nesta área destacamos as Árvores de Regressão (*Breiman, Friedman, Olshen & Stone, 1984*) e as Regras de Regressão (*Cendrowska, 1987*).

Na presente tese, constam algoritmos lógicos de distância e matemáticos, nomeadamente as Redes Neurais (*Basheer & Hajmeer, 2000; Chow & Cho, 2007*), Árvores de Regressão (*Wei-Yin Loh, 2011; De Ville, 2006*), Regras de Regressão (*Witten & Frank, 2005*), algoritmos de Regressão local ou IBL (*Cleveland & Loader, 1996; Atkeson, Moore & Schaal, 1997*), como, por exemplo, KNN (*Han & Kamber, 2001*), Regressão Kernel (*Hastie & Loader, 1993*), Regressão Local linear (*Atkeson, Schaal & Moore, 1997*) e a Regressão local Polinomial (*Katkovnik, 1995*).

ALGORITMOS PARA UMA VARIÁVEL DEPENDENTE	ALGORITMOS PARA MAIS DE UMA VARIÁVEL DEPENDENTE
Regressão Polinomial Regressão Não linear Robusta Regressão Monótona Regressão de Séries de <i>Fourier</i> <i>Wavelet Smoothing</i> Regressão K-Nearest Neighbour Regressão Kernel <i>Smoothing</i> Polinomial Local <i>Smoothing Splines e B - Splines</i> <i>Projection Pursuit</i> , Redes Neurais Métodos Bayesianos <i>Instance Based Learners</i> Árvores / Regras <i>Multivariate Adaptive</i>	Redes Neurais ILP Regressão PLS

Quadro IV.7.1 – Algumas metodologias de análises de Regressão.

Fonte: Elaborado pela autora.

Determinados algoritmos proposicionais de *Data Mining* foram estendidos a bases de dados temporais, relacionais e simbólicas. Nas análises realizadas também constam alguns algoritmos de *Data Mining* simbólicos, por exemplo, as Redes Neurais SOM (*Chow & Cho, 2007; Kohonen, 2001*).

Para alcançarmos os objectivos propostos também será necessário recorrer a metodologias de segmentação dos dados. Na *Figura IV.7.3* enumeram-se algumas metodologias, de Classificação ou Agrupamento de dados (*Berkhin, 2006*) e que estão diferenciadas em métodos paramétricos ou não paramétricos.

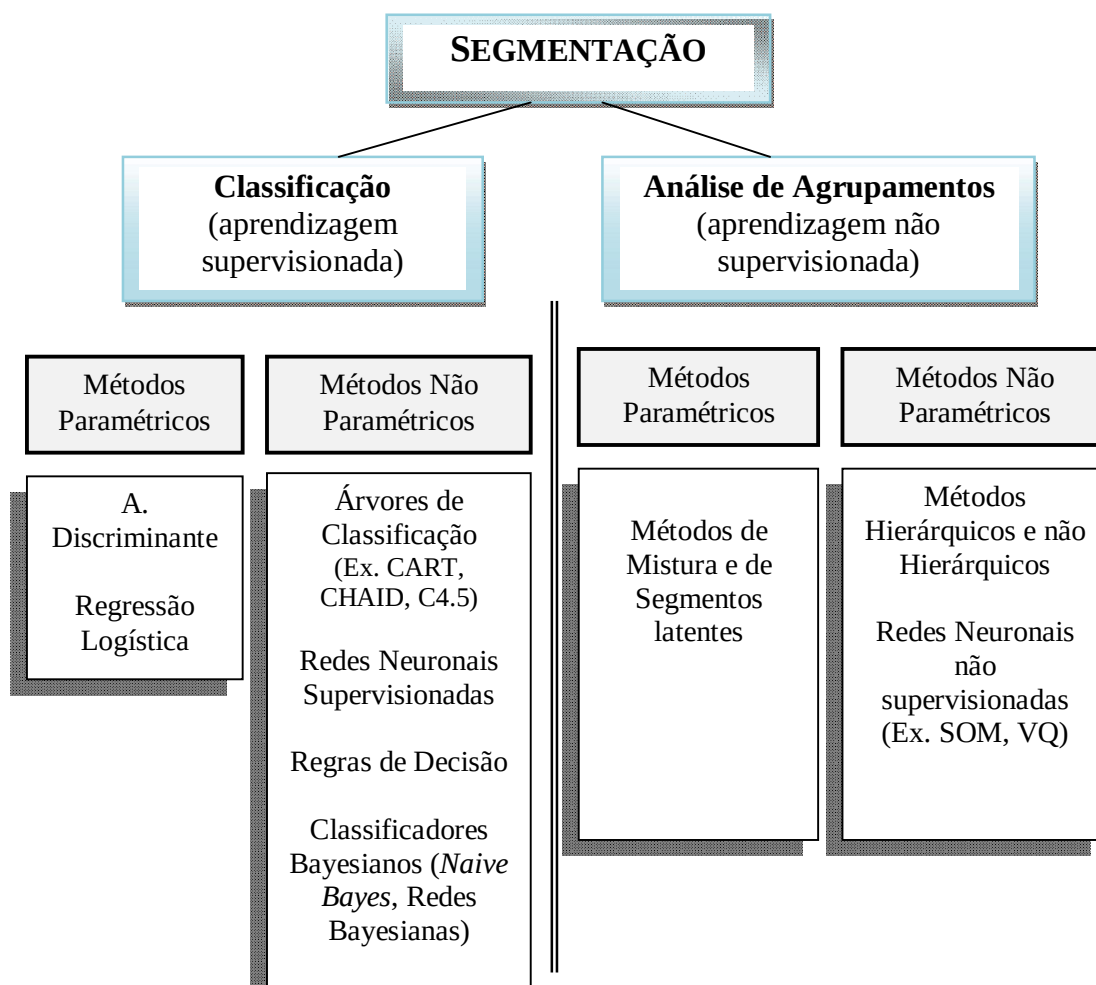


Fig. IV.7.3 – Algumas metodologias de segmentação das bases de dados.
Fonte: Elaborado pela autora.

IV.8. A ESTRUTURA METODOLÓGICA DA TESE

A necessidade de conhecer e prever a evolução dos dados, surge em inúmeros domínios de actividade e os métodos analíticos mais adequados estão, fortemente, dependentes dos contextos. Todos os métodos existentes apresentam limitações. Porém, as abordagens baseadas na aprendizagem automática e no *Data Mining* podem viabilizar, segundo a bibliografia da especialidade, os melhores resultados (Weber, 2002). A selecção do método mais adequado para a previsão da exposição televisiva e dos *ratings* depende marcadamente do contexto, sendo função da disponibilidade de dados,

facilidade de uso e implementação, custos envolvidos e restrições de tempo ou prazos (Geel et al., 2003).

As bases de dados de audimetria nacionais podem apresentar uma evolução significativamente diferente, do que se verifica a nível internacional. Previamente, foram realizadas algumas análises descritivas, que permitiram caracterizar a evolução dos dados. A regularidade da exposição televisiva para diversos níveis de agregação dos dados, as faixas horárias e as tipologias programáticas, entre outros fatores condicionam as abordagens que devem ser adoptadas.

Na área do planeamento de *media* existe um maior interesse por modelos causais de “*ratings*” em detrimento de meras previsões recorrendo às séries temporais de determinados indicadores. Nesta área onde assistimos a uma considerável variabilidade no consumo e nas características de planeamento, para além de outros factores estruturais, os resultados dificilmente se mantêm válidos por um grande horizonte temporal, apresentando os modelos multivariados uma maior relevância prática. Em virtude da existência de sazonalidade e de ciclos económicos no consumo televisivo, devem ser realizadas análises diferenciadas para cada um dos períodos identificados.

No entanto, se o objectivo das análises corresponder apenas à previsão de determinados índices de audiências, podemos recorrer à análise histórica das correspondentes séries temporais, nos períodos em que exista uma maior estabilidade dos dados. Podem ser adoptadas metodologias clássicas de previsão ou algoritmos de *Data Mining* Temporal. Nos períodos onde a exposição é superior e mais variável, será mais questionável a relevância e adequabilidade dos modelos univariados.

Os modelos explicativos de *ratings* diferenciam-se dos modelos de previsão, na medida em que podem incluir variáveis que não são conhecidas nem previsíveis, com uma reduzida margem de erro no momento de realizar a previsão.

O seu objectivo não será apenas o cálculo de valores futuros, mas, sim, a caracterização dos principais condicionantes destes valores e na identificação de interacções que permitam uma melhor percepção do processo subjacente tornando os resultados válidos por um maior horizonte temporal. No entanto, as análises podem permitir identificar algumas variáveis, a ser incluídas em futuros modelos de previsão. Nas faixas horárias a partir do horário nobre, será mais relevante a avaliação das interacções existentes entre as variáveis e devem ser privilegiadas as análises multivariadas.

Atendendo às inúmeras dificuldades que o desenvolvimento dos modelos de *ratings* coloca ao sector, nesta tese também se pretendeu avaliar a potencialidade de diversas áreas metodológicas mais recentes, para a análise das bases de audimetria, em função das necessidades de análise inerentes ao planeamento de *media*, nomeadamente para uma melhor caracterização do processo de selecção programática e para a posterior previsão dos dados.

Nas análises a realizar pretende-se, essencialmente, o desenvolvimento de análises quantitativas e explicativas de determinados índices de audiências programáticos e comerciais, que não negligenciem variáveis fundamentais para a sua posterior previsão. Simultaneamente avalia-se a capacidade preditiva dos modelos e pontualmente também se recorre a metodologias de previsão, verificando a viabilidade de previsão de determinados indicadores a curto prazo e perante horizontes temporais que apresentem uma maior estabilidade.

Podemos adoptar várias perspetivas de análise, quando se pretende a previsão das taxas médias de audiências e a construção de planos de *media*, não sendo viável a implementação de todas as perspetivas neste trabalho. A definição da abordagem metodológica mais adequada implica ter em atenção se será qualitativa versus (vs.) quantitativa; causal vs. não causal; univariada vs. multivariada; enquadrada no *Data Mining*, na Análise Simbólica ou na estatística multivariada; linear vs. não linear e recorrendo a dados agregados para o universo, a segmentos ou aos dados dos ficheiros de base de audimetria.

As abordagens a adoptar dependem fortemente dos contextos, dos objectivos das análises e do grau da tipologia e variabilidade dos dados, que será função das faixas horárias consideradas, composição sociodemográfica da audiência e características da programação envolvida. Sob a hipótese de que será menos difícil prever e caracterizar o comportamento das massas do que o comportamento individual, definido como caótico e imprevisível por inúmeros autores, estabelecem-se previsões e realizam-se análises para conjuntos de telespectadores em função dos seus hábitos de consumo e da exposição televisiva do painel, assim como da estruturação das grelhas, entre muitos outros factores.

Em função dos contextos, a melhor previsão poderá resultar apenas da análise das séries temporais de valores passados e em outros contextos implicar o desenvolvimento de análises explicativas, que promovam a avaliação da influência dos factores estruturais e eventualmente também individuais, ao nível da exposição televisiva e dos indicadores agregados e acumulados. A ocultação da programação e a contra programação ou qualquer outro imprevisto aleatório poderá condicionar a fiabilidade das estimativas realizadas. Segundo *Webster e Litchy (1991)*, no mercado norte-americano consideram-se aceitáveis erros de previsão, que se aproximem dos 10% sobre os valores reais. Para as análises de previsão se adequarem às necessidades do sector, as análises devem incidir nos dados mais recentes e ter em atenção a sua sazonalidade. O meio televisivo encontra-se influenciado por parte de inúmeros factores, que inviabilizam previsões fiáveis a médio ou longo prazo, uma vez que os comportamentos são dinâmicos, evoluindo no decorrer do tempo e também vão existindo sucessivas modificações na constituição do painel.

Uma abordagem poderá corresponder à realização de análises para cada um dos meses ou ciclos económicos, previamente identificados através de análises de segmentação das séries temporais, que considerem a sua variabilidade. Alternativamente os modelos devem incluir a dimensão temporal e sazonal como sugerido em *Patelis et al. (2003)*. Contudo, também devemos tentar tirar o maior partido do conhecimento pré-existente, a respeito da evolução dos dados.

A seguinte *Figura IV.8.1* e a *Figura IV.8.2* representam o conjunto de metodologias utilizadas na investigação empírica realizada neste trabalho.



Fig. IV.8.1 – Conjunto de metodologias utilizadas na tese.
Fonte: Elaborado pela autora.

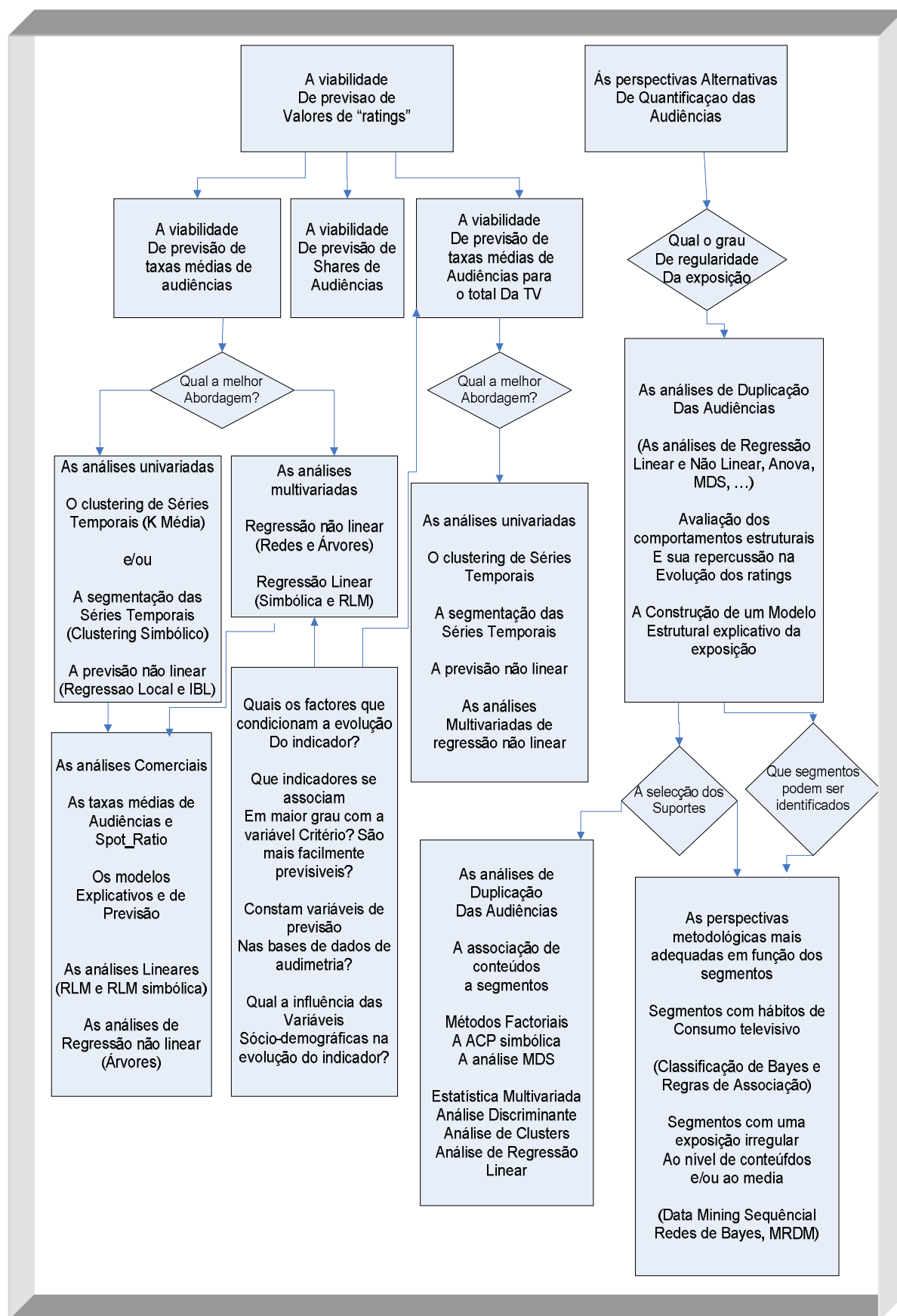


Fig. IV.8.2 – Metodologia conduzida para a pesquisa.

Fonte: Elaborado pela autora.

Consequentemente, atendendo aos factores explicativos considerados como mais relevantes e à necessidade de caracterizar a evolução e variabilidade de alguns indicadores, as análises a realizar visam a avaliação dos seguintes itens, descrevendo-se a sua estrutura do ponto de vista metodológico:

➤ **Análises descritivas das taxas médias de audiências para o total da TV ou TVR%**

Este indicador reflecte a evolução da exposição global dos telespectadores ao media, condicionando as interacções entre os restantes indicadores e variáveis de audimetria. Atualmente persistem dúvidas quanto ao seu grau de independência, face à grelha de programação e em quais são os factores que condicionam a sua evolução e em que extensão. Pretende-se caracterizar a sua evolução para o universo televisivo e para determinados segmentos representados nas bases de dados de audimetria, diferenciando-os ao nível da sua variabilidade. Também se pretende a identificação de alguns dos factores estruturais mais relevantes, quantificando a sua repercussão na evolução desta variável.

Avalia-se o seu grau de previsibilidade, uma vez que esta variável encontra-se cada vez mais influenciada por inúmeros factores, muitos dos quais, dificilmente, quantificáveis e previsíveis. Uma das eventuais opções a que podemos tentar recorrer para tentar prever os valores de *ratings* será por análise da sua associação com *Tvr%* e/ou com os valores de shares, uma vez que resulta do seu produto. Por conseguinte, deve ser avaliada a viabilidade de previsão das taxas médias de audiências para o total da TV.

Segundo alguns autores, para calcular a dimensão da audiência total em determinada faixa horária, dia e mês podemos pressupor que corresponderá ao valor observado para esse período, referente ao ano anterior ou realizar uma análise apenas com base nos seus valores históricos, estudando as tendências observadas nas séries temporais. Consequentemente, será necessário caracterizar a sua evolução e avaliar o seu grau de estabilidade diário, semanal, mensal e interanual.

Para a caracterização da sua evolução recorre-se a metodologias simbólicas de *clustering* e de dissemelhança (Bock & Diday, 2000), assim como à distância euclidiana e ao coeficiente de correlação linear de Pearson, aplicando posteriormente, o algoritmo K Média (Jain, 2010) ou algum algoritmo de *Clustering* Hierárquico (Kaufman & Rousseeuw, 2008).

As publicações internacionais que incidem na análise deste indicador são muito reduzidas e recorrem a metodologias de análise estatística mais tradicionais, como, por exemplo às Regressões trigonométricas, numa década em que existia uma reduzida oferta televisiva. Do ponto de vista metodológico, contrastando com as abordagens identificadas optou-se por metodologias de Análise Simbólica, Previsão (Tavares *et al.*, 1996), *Data Mining* Proposicional e do *Time Series Data Mining (TSDM)* (Esling & Agon, 2012). Seleccionaram-se alguns algoritmos de previsão, *análise de agrupamentos* e segmentação de séries temporais para tentar alcançar melhores resultados do ponto de vista da previsão deste indicador. Também se pretendeu a identificação de faixas horárias comparáveis através da sua segmentação, definindo-se assim, a discretização do fator temporal a incluir posteriormente, nas análises de Regressão multivariadas. O TSDM subdivide-se no “Indexing”, *Clustering*, Classificação, “Novelty Detection”, “Motif Discovery” e no “Rule Discovery”. Avaliou-se em que medida estas áreas podem contribuir para a previsão deste indicador, assim como os métodos simbólicos.

A previsão de *ratings*, em períodos de maiores níveis de audiências, poderá resultar da análise da sua associação linear ou não linear com este indicador. Nesse sentido, realizaram-se algumas análises que contribuem para a previsão de *Tvr%* quando está associado a suportes temporais ou programáticos. Uma vez que a sua evolução não é linear e o horizonte temporal em análise representa um período característico, a nível médio da exposição televisiva, recorreu-se a metodologias de Regressão não paramétrica, como, por exemplo à Regressão *K Nearest Neighbour* (Han & Kamber, 2001; Jain, 2010), Regressão Kernel (Li & Racine, 2007), polinomial (Yang, 2012), local linear (Fox, 2000) e local polinomial (Katzkovnik, 1995), assim como a algumas Redes Neurais.

Para efeito de comparação de resultados também se considerou as análises de Médias Móveis (Tavares *et al.*, 1996). Para a segmentação da emissão diária, em função da sua evolução, foram adoptadas metodologias simbólicas de *clustering*, como, por exemplo o *clustering* divisivo (Gowda & Ravi, 1995; Kim & Billard, 2012) e algumas medidas de dissemelhança para objectos simbólicos booleanos (Esposito, Malerba & Tamma, 2000). Sugerem-se ainda determinadas abordagens metodológicas para a previsão deste indicador, recorrendo aos dados dos ficheiros de base de audimetria, nomeadamente à Regressão linear Múltipla Simbólica (Billard & Diday, 2000; Neto, De Carvalho & Bezerra, 2006) quando se consideram variáveis explicativas da sua evolução ou ao *Data Mining* Temporal (Guimaraes, 2000) para a identificação de padrões nas séries temporais históricas.

A modelização paramétrica consiste em explicar o comportamento de uma variável, a partir do conhecimento que se detém da evolução de outras variáveis. Em geral nas análises de Regressão é pressuposto que as variáveis explicativas estejam fixas por um desenho experimental, ou seja não serem aleatórias. Frequentemente, quando as bases de dados devem ser analisadas recorrendo a metodologias de *Data Mining* tal não sucede, uma vez que as variáveis explicativas se obtêm em contextos observacionais. A função que ilustra a relação entre as variáveis explicativas e a variável resposta, representa a componente determinista e estrutural do modelo. Apresentam-se algumas possibilidades de análise recorrendo a variáveis simbólicas e à Regressão Linear, uma vez que assim podemos representar de forma mais adequada as variáveis, que no decorrer do tempo apresentam variabilidade e diferente distribuição de valores.

A modelização não paramétrica e em particular a que faz uso de métodos núcleo, pode construir modelos mais flexíveis ao ajustar-se aos dados de maneira local. Os métodos paramétricos baseiam-se na hipótese de que os dados seguem determinado modelo matemático, que determina a parte sistemática das observações e a sua componente aleatória. Supõe que os dados resultam da observação de variáveis aleatórias de distribuição conhecida, com excepção da presença de alguns parâmetros cujo valor se desconhece.

Consequentemente, não se adaptam a inúmeros conjuntos de dados, por apresentarem uma estrutura rígida. Por conseguinte, foi necessário recorrer a métodos de ajuste local e a técnicas de aprendizagem supervisionada, como as Redes Neurais e os algoritmos KNN. Em vez da Regressão linear *piecewise* (Ryan & Porth, 2007), que produz discontinuidade nos pontos que definem os vários intervalos sujeitos a interpolação, com os ajustes locais lineares centrados no ponto a prever e considerando as observações na sua vizinhança, podemos ultrapassar esta limitação. Uma vez que a evolução da variável em estudo, apresenta em determinadas faixas horárias uma evolução que se afasta da linearidade e se aproxima da polinomial, também se implementaram as análises locais de Regressão Polinomial. Uma Regressão Polinomial corresponde a uma Regressão Linear, com termos quadráticos e cúbicos. Um caso particular do ajustamento local corresponde ao polinómio de grau zero, designando-se de Regressão Kernel e que também foi adoptada. Os métodos baseados nos vizinhos mais próximos alcançam determinada solução a partir de informação, que é extraída de um conjunto de exemplos previamente existentes.

A desagregação das análises ao nível dos segmentos pretende avaliar quais são as variáveis, que contribuem em maior grau para a variabilidade deste indicador e segmenta-los a este nível. Diferenciam-se os segmentos em função do seu grau de exposição ao meio televisivo e ao nível da sua estabilidade no decorrer da semana. Os segmentos que apresentarem uma maior variabilidade, devem ser alvo de análises posteriores, uma vez que também incutem uma maior variabilidade aos valores de $Tvr\%$ para o universo televisivo. No entanto dificilmente se conseguirá prever de forma “determinista” os padrões de consumo e a exposição global de segmentos, que apresentam uma exposição mais irregular ao meio televisivo e fortemente condicionada por factores estruturais, como, por exemplo a disponibilidade.

Para comparar os segmentos quanto à evolução e grau de estabilidade deste indicador recorreu-se a medidas de dissimilaridade simbólicas e uma vez mais a algoritmos de *clustering* de *Data Mining*, como o K Média (Jain, 2010).

Para avaliar quais são os segmentos que apresentam uma maior consistência nos seus padrões de exposição televisiva no decorrer da semana, recorre-se também à análise em Componentes Principais (Dougherty, 2013).

Pretendeu-se verificar se os melhores resultados para a previsão deste indicador, com base em análises univariadas decorrem da utilização de metodologias de previsão, ou alternativamente de segmentação e *clustering* de séries temporais. Referem-se os contextos em que poderá existir mais-valias em incluir variáveis explicativas ou recorrer aos dados dos ficheiros de base de audimetria e a séries temporais ao segundo, definindo-se também as correspondentes abordagens metodológicas. Finaliza-se com a apresentação de possíveis abordagens metodológicas para a previsão de *Tvr%* associado a valores programáticos e na avaliação da repercussão dos espaços publicitários na sua evolução. Com base nas análises descritivas realizadas foram identificados vários contextos na evolução dos dados que exigem diferenciadas abordagens analíticas, do ponto de vista da caracterização e previsão deste indicador.

Na literatura da especialidade, ainda que a nível internacional, foram escassos os trabalhos identificados que promovem a previsão deste indicador e que o caracterizem de forma mais exaustiva, recorrendo os trabalhos identificados apenas a metodologias de análise estatística tradicionais, assim como a dados para o universo televisivo e incidindo essencialmente nas faixas horárias da noite, que maximizam as audiências televisivas.

As análises descritivas realizadas ao valor da exposição global ao *media*, também permitiram avaliar se em princípio predominam as motivações ritualísticas ou as instrumentais, sendo este assunto desenvolvido em análises posteriores, nomeadamente verificando como o canal e as tipologias programáticas influenciam os resultados e os restantes indicadores de audimetria. A repercussão das variáveis sociodemográficas também será avaliada, uma vez que alguns autores referem que não influenciam, significativamente (Huertas, 2002), o consumo televisivo.

Recorrendo a medidas de dissemelhança simbólicas verificamos, atendendo à variabilidade observada neste indicador, quais são os segmentos que apresentam alguma semelhança relativamente à variável em estudo.

➤ **Análises descritivas aos valores de *shares* de audiências**

Foram realizadas algumas análises exploratórias, essencialmente, de análise gráfica, para proceder à caracterização da evolução dos valores de *shares*. Este índice é função da exposição global dos telespectadores ao meio televisivo, mas, também, de eventuais preferências e da sua tendência para desenvolverem comportamentos estruturais, que derivam da estruturação da grelha de programas, entre muitas outras variáveis estruturais e individuais. Os telespectadores podem apresentar diferentes comportamentos ao nível da lealdade programática, tipológica e ao canal, sendo esses comportamentos estruturais determinantes, para conseguir explicar e prever a evolução deste indicador.

Nesta tese, constam algumas análises, meramente exploratórias, aos valores dos *shares* programáticos e temporais para avaliar o seu grau de estabilidade. A elevada variabilidade deste indicador no decorrer da maior parte da emissão diária, dificulta a sua utilização em análises de previsão, e em especial nos períodos de maiores audiências. Os *shares* medem como a audiência total se distribui entre os vários canais, os quais nestes períodos optam por transmitir os melhores programas e recorrem a estratégias de programação, frequentemente, competitivas. A informação que consta das bases de audimetria, ou que poderá ser calculada a partir dela, dificilmente deverá apresentar uma significativa capacidade preditiva desta variável. Para compreender a sua evolução foram posteriormente realizadas, algumas análises aos indicadores acumulados de audimetria, que pretenderam avaliar em que medida estão presentes os comportamentos estruturais de exposição televisiva no painel de audimetria nacional.

➤ **Análises descritivas aos valores de *Ratings* ou Taxas Médias de Audiências**

O *rating* corresponde ao indicador mais relevante para o planeamento de *media*. O seu valor previsto poderá corresponder à evolução histórica desta variável para idênticos períodos horários, sendo necessário realizar análises univariadas por aplicação de metodologias de análise de previsão das ST, mas, também, da área da segmentação e *clustering*.

Provavelmente, os períodos horários que apresentem uma menor exposição televisiva, também podem minimizar a sua variabilidade viabilizando esta opção. Para a segmentação e previsão deste indicador recorreu-se às abordagens metodológicas anteriormente referidas, para a análise das taxas médias de audiências para o total da TV. Do ponto de vista analítico contrastando com as publicações disponíveis recorre-se ao *clustering* (Prakash & Li, 2011), segmentação (Lemire, 2006) e a algoritmos de previsão não linear de séries temporais (Cunningham & Delany, 2007).

Pretende-se, também, caracterizar a sua evolução e avaliar quais são os factores que o condicionam em maior grau, e que posteriormente podem ser incluídos em modelos de previsão. Diferenciam-se os contextos em que as análises de previsão devem incidir em análises univariadas ou multivariadas dos indicadores de audimetria ou recorrer aos dados dos ficheiros de base de audimetria. Também se avalia a repercussão dos eventos desportivos na sua evolução, uma vez que estas situações podem requerer que sejam adoptadas diferentes abordagens analíticas. Nestas situações devemos tentar identificar regularidades ao nível das alterações que se verificam nas séries temporais multivariadas de valores de *ratings*, referentes aos vários canais generalistas. Análises mais pormenorizadas a este nível não são realizadas, uma vez que é mais relevante a análise dos comportamentos estruturais e das audiências comerciais, para se poder apresentar uma proposta no âmbito dos modelos de exposição.

➤ **Análises descritivas aos Indicadores Acumulados de Audiências Televisivas**

Nos períodos de maximização da exposição televisiva onde se observa uma maior variabilidade nos dados, os modelos preditivos explicativos são mais relevantes, assumindo os comportamentos de duplicação de audiências uma maior relevância, assim como a interação existente entre a vertente programática e comercial.

Constam algumas análises realizadas aos indicadores acumulados, que permitem verificar qual é o grau de actividade e regularidade inerente à exposição e selecção televisiva. Contrastando com a literatura da especialidade, as abordagens metodológicas também têm origem em áreas mais recentes de análise estatística de dados.

➤ Análise dos índices de lealdade programática e horária, ou seja dos índices de duplicação de audiências entre episódios e transmissões de um mesmo programa, ou considerando os suportes horários.

A quantificação da duplicação das audiências poderá eventualmente contribuir, para a construção dos modelos de exposição. De acordo com *Fernandes (2001)*, a análise da duplicação das audiências, insere-se nas abordagens que visam a previsão das audiências, uma vez que caracterizam a tendência que se observa nos telespectadores de determinado suporte, para também estarem presente na audiência de um suporte posterior. Estas análises também condicionam a selecção dos suportes a integrar nos planos de *media* e permitem avaliar o grau de regularidade da exposição e interpretar qualitativamente a evolução dos *ratings*.

Existindo regularidade na exposição, posteriormente referem-se algumas abordagens metodológicas, que pretendem o cálculo da probabilidade de determinado telespectador assistir a sucessivos episódios de determinada série programática, considerando variáveis explicativas destes comportamentos ou não.

Os métodos Bayesianos (*Becker, Kohavi & Sommerfield, 2001*) permitem a avaliação da repercussão de factores estruturais a este nível e as médias móveis exponenciais (*Cannon, 1996*) viabilizam a previsão das suas futuras seleções recorrendo apenas ao histórico de seleções para cada telespectador. Tradicionalmente na literatura da especialidade dos modelos de duplicação das audiências, apenas são consideradas análises de Regressão linear Múltipla (*Cameron & Pravin, 2005; Kleinbaum, Kupper, Muller & Nizam, 1998*), de Estatística Descritiva e Análises de Variância (ANOVA) (*Keppel & Wickens, 2004*) e associadas essencialmente ao horizonte temporal que maximiza a exposição televisiva.

Os trabalhos inicialmente publicados nesta área medem a duplicação das audiências, considerando a distribuição de frequências cumulativas para um conjunto de programas. Em *Zubayr (1999)*, o autor vai mais além e considera como variável dependente as audiências conjuntas entre os programas que são transmitidos de forma consecutiva, ou seja a percentagem da audiência que contactou com ambos os programas ($P(i \wedge j)$ com i e j sequenciais). A terminologia leva-nos a pressupor, que será suficiente um segundo de exposição para ser considerado como telespectador em cada par de programas considerado.

Nas análises realizadas, a variável dependente corresponde a taxas de audiências condicionadas ($P(i / j)$) e os suportes não correspondem necessariamente a programas adjacentes. Para o seu cálculo consideram-se os telespectadores que assistiram no mínimo a cinquenta por cento da duração de cada um dos programas, incluídos em cada par de observações e também exposições mais breves. Considera-se a totalidade das faixas horárias e maiores desfasamentos temporais entre os episódios, para além de serem apresentadas metodologias que também permitem caraterizar o perfil dos telespectadores, que apresentam maior ou menor lealdade a determinados programas, como, por exemplo a análise binomial Logit (*Kleinbaum, 2010*).

As análises *Binomial Logit* surgem na bibliografia da análise estatística de dados, quando se pretende a análise de comportamentos de lealdade face a determinados produtos ou serviços, permitindo definir as características dos consumidores, que apresentam determinado perfil. Sugere-se para o contexto televisivo, quais são algumas das variáveis explicativas, que podem apresentar uma maior relevância e a sua codificação, considerando os resultados das análises realizadas nos capítulos anteriores.

Recorrendo a análises de correlação e ANOVA (Keppel & Wickens, 2004) avalia-se a repercussão de algumas variáveis estruturais de planeamento, na evolução dos índices de lealdade programática, nomeadamente a influência do canal, das tipologias programáticas, desempenho dos programas, faixas horárias, tipo de programa e o número de transmissões por semana. Comparam-se os resultados alcançados com as conclusões apresentadas a nível internacional.

➤ Análise da duplicação de audiências entre programas directamente adjacentes ou não; das decisões de ligar e desligar o televisor e de fluxos de audiências entre canais (Efeitos de Herança).

Conseguir favorecer os índices de duplicação e minimizar os fluxos para os canais concorrentes, permite aos canais poder maximizar e apresentar em determinados períodos uma maior estabilidade nos valores dos seus índices de audiências, facilitando a previsão de valores de *ratings* (Apêndice F.4). A repercussão de alguns factores estruturais, ao nível dos efeitos de herança entre programas directamente adjacentes é avaliada na literatura da especialidade, recorrendo a análises de Regressão linear Múltipla e incidem nos períodos de maiores audiências. Neste trabalho privilegiam-se os métodos não lineares do *Data Mining* Proposicional (Wei_Yin Loh, 2011), sugerindo-se abordagens alternativas ao nível do *Data Mining* Temporal e em especial do “*Temporal Reasoning*”.

Inclui-se um conjunto de variáveis explicativas estruturais muito mais abrangente, face ao que consta dos trabalhos internacionais publicados, que regra geral apenas consideram os *ratings* dos programas envolvidos, para além das suas tipologias e o número de programas que tem início na sua interface. Em vez de restringir as análises aos dados programáticos adjacentes, também se considera os dados referentes a períodos horários e a programas não adjacentes, ou seja os dados do primeiro e segundo nível programático das bases de dados *Telereport*. A variável dependente representa a duplicação primária ou taxa de audiência condicionada. Este indicador surge com uma menor frequência na literatura da especialidade, ainda que seja o mais adequado para a avaliação e caracterização destes comportamentos de consumo televisivo. Também foram realizadas análises com base nos valores de *ratings* de programas adjacentes, sendo esta abordagem mais relevante do ponto de vista da previsão dos *ratings*. Pretendeu-se também contribuir para esclarecer algumas questões que constam da publicação mais recente que foi identificada (*Webster, 2006*). Nomeadamente porque observações que não partilham um forte número de telespectadores apresentam por vezes, valores muito próximos de valores de *ratings*.

Nas bases de dados consideram-se os fluxos de audiências entre os canais, o que permite avaliar se a duplicação de audiências se encontra condicionada em maior grau pelos fluxos de telespectadores para os canais concorrentes ou pelas decisões de ligar e desligar o televisor.

Do ponto de vista metodológico a utilização da análise de Correspondências (*Greenacre & Blasius, 2006; Hair, Anderson, Tatham & Black, 1998*) e da análise MDS (*Borg & Groenen, 2005; Wickelmaier, 2003*) permite a diferenciação dos períodos horários e programas, em função do seu grau de duplicação de audiências. Identificamos com base em valores médios, quais são os períodos horários que partilham um maior número de telespectadores, o que também permite identificar segmentos de público, em função da faixas horárias.

Posteriormente agregaram-se às bases de dados de duplicação de audiências primárias, as variáveis que traduzem os fluxos de audiências entre os canais generalistas e face aos canais por cabo, assim como as variáveis que quantificam as decisões de ligar ou desligar o televisor.

Utilizam-se as Árvores de Regressão CART (*Breiman et al., 1984; Wei-Yin Loh, 2011*) e CHAID (*Abdi, 2007, Kass, 1980; Magidson, 1993*) que viabilizam uma análise de Regressão não linear para avaliar as interações entre os vários factores e realizam uma previsão não linear. Não foi possível identificar, na literatura da especialidade, análises que avaliem a repercussão dos fluxos de audiências entre os canais e das decisões de ligar e desligar o TV na evolução dos índices de duplicação. Também se sugere uma abordagem analítica que permita tentar quantificar a duplicação de audiências.

➤ Análise de diversos indicadores sociodemográficos de fidelidade, adesão, atractividade e afinidade que caracterizam a exposição dos diversos segmentos e permitem avaliar se existem comportamentos de lealdade a canais, tipologias e programas.

Estas análises pretendem definir o perfil de audiências dos diversos suportes; verificar em que medida as variáveis sociodemográficas constituem variáveis discriminantes da exposição televisiva; caracterizar o grau de envolvimento do painel com a grelha de programas e promover a segmentação das audiências televisivas.

Para as análises de segmentação são sugeridas determinadas abordagens metodológicas e diferenciadas em função da tipologia de dados em que incidem, isto é, quando se recorre a dados agregados ou aos dados dos ficheiros base de audimetria. A associação de suportes a segmentos permite reduzir a dimensão da amostra, na qual irão incidir as análises posteriores, que viabilizam a construção das campanhas publicitárias. Uma maior fidelização do público face à grelha de programas contribui para uma maior estabilidade dos seus índices de audiências e para uma menor repercussão da programação concorrente, influenciando os efeitos de herança e os valores de *shares*.

Os índices de fidelidade permitem avaliar o grau de envolvimento dos telespectadores com os programas transmitidos. Correspondendo a um indicador percentual da densidade dos consumos, representa a qualidade da audiência média. Quanto mais elevado for o seu valor, maior será a capacidade de um programa para conseguir manter os telespectadores no seu decorrer. Os restantes índices sociodemográficos permitem a identificação de segmentos de telespectadores para determinados programas e a associação das tipologia programáticas e dos canais aos mesmos. Permitem avaliar se existem comportamentos de lealdade, contribuindo assim, para a futura quantificação da exposição televisiva. Pretende-se identificar quais são os níveis das diversas variáveis de segmentação, que apresentam padrões de exposição ao *media* mais semelhantes.

Com a análise das bases de dados de audimetria, apenas podemos avaliar a influência das variáveis sociodemográficas. Mas, do ponto de vista prático, também, são as mais utilizadas em análises de segmentação, uma vez que são facilmente identificadas, mensuráveis e podem ser calculadas com um baixo custo, apresentando-se em princípio como um bom fator, que pode mediar a previsão das escolhas realizadas.

Do ponto de vista metodológico, as abordagens de associação de suportes a segmentos diferenciam-se em função das tipologias programáticas em análise.

➤ Para a análise de programas que apresentam apenas uma transmissão recorre-se a metodologias de estatística multivariada. As análises MDS (*Borg & Groenen, 2005*) permitem associar programas a segmentos com base em medidas de dissimilaridade, recorrendo também à Regressão linear (*Kleinbaum et al., 1998*) para a interpretação dos eixos factoriais. Alternativamente recorre-se a algoritmos de *clustering* hierárquicos ou de *Data Mining* e à análise Discriminante linear (*Duda, Hart & Stork, 2000*) para tentar identificar quais são as variáveis de segmentação que permitem discriminar os agrupamentos de programas identificados. A análise Discriminante linear contribuiu para a identificação de Regras de atribuição de observações, a cada uma das classes ou grupos pré-estabelecidos ou definido com base em análises de agrupamento dos dados.

➤ Para a análise dos programas transmitidos numa sequência de episódios, independentemente de lhes estar, ou não, associada uma sequência narrativa, adoptaram-se algoritmos simbólicos de *clustering* (Gowda & Ravi, 1995; Kim & Billard, 2012), factoriais (Lauro & Palumbo, 2000) e de MDS (Groenen, Winsberg, Rodriguez & Diday, 2005, 2006). Na área do *Data Mining* privilegiaram-se alguns algoritmos de *clustering* (Berkhin, 2006) e as Regras de Associação ou Sequenciais (Xin et al., 2007). As Regras de Associação (Zaki, 2001) viabilizam a identificação de qual é o conjunto de programas, ou intervalos de tempo, mais seleccionado em cada sessão televisiva por determinado segmento ou telespectador, com base em medidas de confiança e cobertura. No DM sequencial é relevante a ordem em que os suportes são seleccionados, e permite descobrir associações entre suportes visionados em diversas sessões televisivas.

Nas análises realizadas às variáveis sociodemográficas de *ratings* pretendeu-se segmenta-las em função da evolução dos indicadores, verificando quais são os segmentos que apresentam uma evolução mais próxima entre si e face ao universo televisivo. Permitiu identificar quais são os agrupamentos de telespectadores responsáveis, em maior grau, pela variabilidade observada nos valores dos *ratings* para o universo televisivo.

As metodologias simbólicas de *clustering*, como, por exemplo as Redes Neurais SOM (Kohonen, 2001) ou as medidas de dissemelhança são as perspectivas de análise mais adequadas. As Redes Neurais de aprendizagem não supervisionada caracterizam-se por descobrir, unicamente a partir dos dados de entrada, modelos e características que são considerados significativos. As Redes Neurais SOM são autoorganizativas e, conseqüentemente, devem actuar perante alguns aspectos do conjunto de entrada, como, por exemplo a existência de redundância ou a existência de grupos no conjunto de dados e preservam a tipologia.

Na literatura da especialidade, não foram identificadas análises aos índices sociodemográficos de audimetria. No âmbito dos trabalhos que visam avaliar os comportamentos de lealdade, foram identificadas algumas publicações com dados de audimetria Alemães e Britânicos que correlacionam o número de programas seleccionados de determinada tipologia programática com as restantes tipologias ou ao nível dos canais seleccionados, utilizando metodologias mais tradicionais, como a Regressão linear e medidas de Estatística Descritiva.

Em *Kim (2002)* constam análises de segmentação do painel de audimetria Coreano, recorrendo a um índice de co-ocorrência entre programas e à análise de *clusters*. Nesta tese são sugeridas algumas abordagens metodológicas alternativas. Quando os anunciantes utilizam este *media* como um veículo de comunicação e divulgam os seus produtos e serviços com base em estratégias direccionadas para determinados nichos de mercado, é necessário conhecer a composição sociodemográfica da audiência de cada um dos programas e canais, assim como a predisposição do público para desenvolver comportamento de lealdade. Quanto maior for o grau de envolvimento do painel com a grelha de programas, maior deverá ser também a atenção que é direccionada aos seus espaços comerciais.

Os anunciantes pretendem que os seus produtos estejam associados aos programas que maximizam a exposição ao público-alvo das campanhas e aos conteúdos que a audiência possa considerar como sendo mais interessantes, em função da oferta disponível em determinado momento. Os contratos de aquisição de tempo publicitário são celebrados definindo-se o preço, datas e hora de transmissão das diversas inserções. Deve ser garantido um valor mínimo da exposição ao segmento sociodemográfico pretendido ou para o universo televisivo. Os canais garantem um valor global de audiências acumuladas que terá que ser alcançado independentemente de alguns programas poderem apresentar valores superiores ou inferiores ao esperado. Consequentemente, a previsão dos *ratings* sociodemográficos e o conhecimento do perfil das audiências é determinante neste processo.

➤ **Análises Multivariadas de índices de audiências programáticas**

Realizaram-se algumas análises multivariadas, que procuram inter-relacionar e avaliar as associações existentes entre as variáveis de audimetria. Uma vez avaliados de forma independente, alguns dos fatores que podem condicionar a evolução dos principais índices de audiências, foi necessário o desenvolvimento de modelos explicativos multivariados, que também permitam a análise das respetivas interdependências, por aplicação de metodologias lineares e não lineares de Regressão. Avaliou-se a capacidade explicativa das análises multivariadas e identificaram-se as variáveis mais relevantes do ponto de vista da previsão.

Na literatura da especialidade apenas foi possível identificar algumas análises que recorrem à Regressão linear e que incluem como variáveis explicativas as tipologias programáticas e as variáveis temporais, para além de um reduzido número de indicadores. Nas análises realizadas incluiu-se a totalidade dos indicadores que constam das bases de audimetria, para além de outras variáveis que foram calculadas para caracterizar a estruturação da grelha. As metodologias lineares adotadas foram Regressão linear Múltipla (*Kleinbaum et al., 1998*) e o equivalente algoritmo simbólico (*Billard & Diday, 2000; Neto, De Carvalho & Bezerra, 2006*).

As metodologias não lineares utilizadas foram alguns algoritmos de *Data Mining* Proposicional, como, por exemplo, as Redes Neurais *Multilayer Perceptron* (MLP) (*Khoshhal & Mokaram, 2012*) e as Redes Neurais *Radial Basis Function* (RBF), as Árvores de Regressão CART (*Breiman et al., 1984*) e CHAID (*Magidson, 1994*) e Regras de Regressão. Os dados incidiram na totalidade da emissão diária, uma vez mais contrastando com a literatura da especialidade.

A Regressão linear permite avaliar o grau de linearidade das relações entre as diversas variáveis explicativas e serve de base de comparação, para avaliar o diferencial da capacidade preditiva quando se adotam metodologias não lineares ou os algoritmos simbólicos.

Quando existem problemas de colinealidade, e nas análises a realizar se verificou esta situação em determinados contextos, podemos seleccionar as variáveis que não estão correlacionadas ou realizar a Regressão sobre uma transformação não correlacionada, das variáveis originais recorrendo à Regressão em Componentes Principais.

Atendendo aos problemas já mencionados, relativamente aos métodos paramétricos que se baseiam na hipótese de que os dados seguem determinado modelo matemático, que determina a parte sistemática das observações e a sua componente aleatória, também foi necessário recorrer a metodologias não paramétricas. Aplicaram-se algumas técnicas de aprendizagem supervisionada, como as Redes Neurais MLP (*Tang, 2007*) e RBF (*Chow & Cho, 2007*) e as Árvores de Regressão que permitem realizar uma previsão não linear e avaliar as interacções entre as diversas variáveis explicativas.

As Árvores de Regressão permitem a identificação de modelos que são compreensíveis e proposicionais. Representam as regularidades que existem nos dados de forma simbólica, através de um conjunto de condições, o que contrasta com os resultados das Redes Neurais.

As Árvores proporcionam modelos que são inteligíveis, inclusive para sistemas semiautomáticos que processem Regras. Uma das suas vantagens consiste em que na sua forma geral as opções possíveis a partir de determinada condição são excluentes. Isto permite a análise de uma situação e que seja alcançada uma única decisão.

As bases de dados que não são linearmente separáveis, podem ser modeladas com a utilização das Redes que apresentam uma forma de cascata. As Redes Neurais MLP representam sistemas conexionistas com grande capacidade de generalização, viabilizando a identificação de padrões nos dados e a sua previsão não linear. Apresentam uma grande flexibilidade, adaptando-se a novas situações sem necessidade de serem novamente programadas. Trabalham com informação que se encontra incompleta e quando existe ruído nos dados. As Redes Neurais MLP foram treinadas segundo o algoritmo de retropagação de erros.

A segunda tipologia de Rede Neuronal adotada foi as funções de base radial, onde as funções de cálculo correspondem a funções gaussianas, em substituição das funções de activação das Redes Neurais MLP.

Ambas as Redes Neurais são excelentes aproximadores universais, ou seja modelam as funções contínuas com uma grande proximidade, mas apresentam diferenças a diversos níveis. Todas as neuronas das Redes Neurais MLP calculam uma mesma função sobre os valores das variáveis de entradas, que são, previamente, ponderados e proporcionam um cálculo global. As neuronas ocultas das Redes Neurais RBF realizam uma projecção não linear, mas a sua capa de saída será sempre linear e é calculado o somatório das saídas locais. A não linearidade das MLP, geralmente, segue uma função monótona, mas nas Redes Neurais RBF utiliza-se uma função que decai radialmente. As Redes Neurais MLP apresentam uma maior capacidade para o cálculo de soluções, em regiões do espaço de entrada onde existe uma maior escassez de dados no conjunto de observações de treino. Quando se requer resultados muito precisos em todo o espaço de treino, será necessário a utilização de inúmeras funções de base radial, ou seja de neuronas ocultas, sendo esta uma das suas limitações. Porém, devido à natureza local dos modelos, as Redes Neurais RBF são menos sensíveis à ordem em que se apresentam e processam os dados. Nas Redes Neurais MLP devemos retropropagar os erros para conseguir alterá-los progressivamente, enquanto nas RBF não será necessário e, logo, são mais fáceis de treinar.

As Redes Neurais apresentam a desvantagem de apresentarem um funcionamento opaco, que não permite compreender a evolução dos dados. O algoritmo CHAID, assim como as Regras de Regressão, permitem a avaliação das interdependências existentes entre as variáveis e a sua hierarquização, em função da sua capacidade explicativa e preditiva. Contribuem para a posterior construção de modelos lineares, onde as variáveis simbólicas podem viabilizar uma análise mais adequada dos dados de audimetria, permitindo alcançar uma maior percentagem de variância explicada da evolução dos dados.

➤ **Análises multivariadas de índices de audiências comerciais: *Ratings* (*Rat_Spot%* e *Rat_Break%*) e *Spot_Ratio%***

Foram realizadas análises aos *ratings* comerciais, considerando inserções publicitárias (*Rat_Spot%*) e espaços comerciais no seu todo (*Rat_Break%*), e também ao indicador que os inter-relaciona com a vertente programática (*Spot_Ratio%*). As metodologias adoptadas coincidem com as que foram seleccionadas para a análise dos valores dos *ratings* programáticos.

Considerou-se a Regressão linear múltipla clássica e a Regressão simbólica e do ponto de vista não linear as Árvores de Regressão CART e CHAID. As análises inserem-se na perspetiva da investigação presencial, avaliando-se a repercussão de um vasto conjunto de variáveis explicativas na evolução dos *ratings*, a maioria das quais não consideradas na literatura da especialidade e por aplicação de metodologias simbólicas e não lineares que também não foram identificadas.

Do ponto de vista metodológico consta uma exaustiva comparação, das vantagens e desvantagens das metodologias que foram seleccionadas. Não foram identificadas publicações que visem a previsão dos indicadores comerciais e são muito escassas ao nível dos modelos explicativos.

➤ **Proposta metodológica para o planeamento de *Media*: A construção dos modelos de exposição**

A presente tese conclui com a identificação de algumas perspetivas de análise, que podem conduzir a melhores resultados na área dos modelos de exposição. Avaliou-se o contributo do DM Sequencial, DM Multirrelacional, DM Temporal, metodologias Bayesianas e algoritmos de previsão simbólicos. Em função dos resultados das análises exploratórias e descritivas realizadas, são avaliadas as melhores perspetivas em função dos contextos.

➤ **A validação dos resultados das Metodologias de *Data Mining***

Para a avaliação da capacidade explicativa e preditiva das metodologias de Regressão de DM recorreu-se a vários métodos de validação (Arlot, 2010). O método “*hold out*” consiste na divisão do conjunto de dados em duas amostras independentes: amostra de treino e amostra de teste. A amostra de treino corresponde a 70% das observações e a amostra de teste a 30% das observações. Nas análises de Redes Neurais recorrendo ao *software* SAS Enterprise Miner versão 6.2 consideraram-se três amostras de dados: amostra de treino com 40% das observações utilizada para construir os modelos, amostra de teste com 30% das observações que se destina à avaliação do modelo ajustado e amostra de validação com 30% das observações utilizada para a selecção dos melhores modelos, através da avaliação do seu desempenho num conjunto de dados que não foi utilizados para o seu desenvolvimento.

Alternativamente, consideraram-se determinados métodos de reamostragem, como a validação cruzada “*K-fold*” e o “*Bootstrap*”. Na validação cruzada a amostra é dividida em k subamostras, sendo $k-1$ conjuntos de dados utilizados como base de treino e a restante como amostra de teste. O processo é repetido K vezes e no final os erros são ponderados para se calcular um valor global do erro, tendo sido seleccionado um valor de K no valor de 10. Não foi utilizada a validação cruzada “*leave one out*”, onde a amostra de treino exclui apenas uma observação em sucessivas iterações e, logo, são validados n modelos a partir de $n-1$ registos na base de dados.

Foi privilegiada a opção de subdivisão dos dados em amostra de treino e teste, mas seleccionaram-se os métodos alternativos nos contextos de menor dimensão da amostra, porque inviabilizaria a primeira opção. Para a validação das metodologias de Regressão linear também foi utilizada uma amostra de teste, com a mesma percentagem de observações.

IV.9. OS SISTEMAS INFORMÁTICOS UTILIZADOS NA REALIZAÇÃO DAS ANÁLISES

Para a realização das análises estatísticas realizadas recorreu-se aos seguintes *softwares* de análise estatística de dados e de *Data Mining*:

- *Software* STATISTICA Advanced versão 6.0;

As análises de estatística multivariada foram realizadas recorrendo ao *software* STATISTICA: Regressão linear Múltipla, MDS, Discriminante Linear, ANOVA, análise de *Clusters*, análise de Componentes Principais;

- *Software* IBM SPSS *Answer Tree* versão 12.01: Árvores CART e CHAID de Regressão;
- *Software* SAS/STAT versão 9.2;
- *Software* SAS Enterprise Miner versão 6.2 (SAS 9.2 release): Redes Neurais MLP, Redes Neurais RBF, K Média, Regras de Associação e Sequenciais;
- *Software* SODAS versão 2 (*Diday & Noirhomme-Fraiture, 2008*): Análises de *Clustering* e de Dissemelhança, Factoriais e de MDS, Redes SOM;
- Microsoft Office ACCESS versão 2003 e 2010 para Windows XP;
- *Software* WEKA (*Hall, Eibe, Holmes, Pfahringer, Reutemann & Witten, 2009*)

CAPÍTULO V

APRESENTAÇÃO E DISCUSSÃO DE RESULTADOS

Neste capítulo pretende-se a avaliação das questões apresentadas no âmbito dos modelos de “*ratings*” e que são essenciais para a actividade de planeamento de *media*. A construção das melhores perspectivas de quantificação das audiências televisivas depende, inevitavelmente, da caracterização e compreensão da evolução dos indicadores agregados e acumulados de audimetria, e, conseqüentemente, de uma maior percepção a respeito dos factores que condicionam a exposição televisiva e, também, a evolução dos índices de audiências.

Apenas consta a discussão dos principais resultados da investigação empírica e, em função dos quais, no capítulo seguinte, se apresentará uma proposta síntese metodológica que viabilize a construção de determinados modelos que são necessários ao investimento publicitário. A exposição televisiva poderá ser quantificada através de diversos indicadores, existindo múltiplas perspectivas de análise, que podem viabilizar a construção das campanhas publicitárias e que são função do grau de regularidade dos indicadores e dos comportamentos que os originam.

Apresentam-se algumas conclusões substântivas e metodológicas relacionadas com os vários contextos que condicionam em maior grau a aquisição do tempo comercial, como, por exemplo, relativamente aos índices de audiências que quantificam a duplicação de audiências entre os suportes publicitários ou que são necessários para o cálculo de uma medida da pressão total de uma campanha publicitária.

V.1. INTRODUÇÃO

A presente tese pretendeu abordar uma questão que apresenta uma grande relevância prática. Os modelos de exposição, do ponto de vista tático, são essenciais para a estratégia de espaços e tempos, compra e verificação da comunicação, e a nível estratégico contribuem para a selecção do conjunto de meios, pressão e sazonalidade da campanha, assim como para a construção das grelhas e selecção dos conteúdos a transmitir.

Pretendemos identificar quais são as perspectivas de análise que podem cotribuir para uma selecção mais adequada dos suportes publicitários e/ou quantificação dos valores das suas audiências futuras. As áreas metodológicas em avaliação resultam das características das fontes de dados necessárias ao planeamento de *media*, assim como das características dos comportamentos de consumo televisivo e dos objectivos das análises propostas. Na perspectiva adoptada, destacam-se os instrumentos analíticos, para além da avaliação de algumas influências estruturais na evolução dos “*ratings*”, uma vez que condicionam a identificação das perspectivas metodológicas, que podem conduzir aos resultados mais satisfatórios na área da construção dos planos de *media*.

Corresponde a um trabalho de cariz analítico, no âmbito da investigação quantitativa e alicerçado principalmente nas teorias estruturais explicativas do consumo televisivo, que promovem a análise da exposição televisiva e o desenvolvimento de análises explicativas e de previsão dos “*ratings*”, que por sua vez condicionam a sustentabilidade económica do sector.

A perspectiva estrutural tem sido considerada, no passado e a nível internacional, como a que poderá apresentar uma maior capacidade de previsão dos padrões de consumo televisivo da população.

A integração da Estatística e do *Data Mining* poderá conduzir a melhores resultados, uma vez que se diferenciam relativamente às suas condições de aplicabilidade (Branco, 2010, p.43): “A estatística e DM, embora distintas, partilham objectivos e métodos, a nossa convicção é que a cooperação entre os especialistas das duas áreas proporciona um avanço mais rigoroso, mais seguro e mais rápido do processo geral da análise de dados.”

Consideraram-se múltiplos contextos descritivos da evolução dos dados e alguns de previsão, e recorre-se à única fonte de dados que viabiliza uma análise diacrónica da audiência, constituindo o principal instrumento de previsão e de planificação que existe na actualidade. Do ponto de vista da abordagem estrutural de investigação das audiências, os dados constituídos por “*ratings*” representam uma das principais fontes de dados, sendo a mais utilizada na investigação académica Anglo-Saxónica.

As análises explicativas realizadas encontram-se no campo oposto ao da investigação de audiências que se fundamenta nos pressupostos da perspectiva de análise das gratificações e usos do *media* e, que se associa em maior grau às Ciências Sociais, nomeadamente à Psicologia, por análise das motivações e interesses dos telespectadores. Na realidade, a exposição televisiva encontra-se em maior grau dependente das motivações ritualísticas dos telespectadores (Cooper & Tang, 2009; Wonneberger et al., 2012) que visam o preenchimento de um tempo de lazer ou “vazio” e, fortemente, condicionado pelas estruturas. As perspectivas de análise da exposição televisiva integrativas teóricas e empíricas, considerando variáveis estruturais e individuais, são muito escassas na bibliografia da especialidade.

Nas primeiras questões em avaliação pretende-se, essencialmente, uma maior percepção sobre como o painel interage com o *media*, através da análise dos índices de duplicação das audiências e da construção de um modelo de caracterização da sua exposição, e em função do qual se avaliará a viabilidade de serem realizadas análises de previsão e a estrutura metodológica mais adequada (Questão 2).

Também se pretende a identificação dos factores que condicionam a evolução dos indicadores, que resultam dos comportamentos do painel de audimetria e que são essenciais para a construção dos planos de *media* ou para a aquisição de suportes publicitários no “Spot Market”, ou seja, os *ratings* que, por sua vez, dependem das taxas médias de audiências para o total da TV e dos *shares* de audiências, entre outros indicadores (*Questão 1*).

As análises exploratórias realizadas e que constam do *Apêndice F.1* tiveram como objectivo caracterizar os indicadores em avaliação na *Questão 1* e os *Apêndices F.3, F.4* e *F.5* referem-se às análises de duplicação das audiências que permitem a avaliação da regularidade da exposição, mas, também, contribuem para a selecção dos suportes atendendo ao princípio da frequência óptima e estrutural. A análise dos indicadores acumulados, também, contribui para a construção de um modelo explicativo da exposição, que põe em evidência os múltiplos factores estruturais que condicionam o consumo televisivo e a viabilidade da sua previsão. Os indicadores acumulados permitem uma interpretação da evolução de alguns indicadores agregados, como, por exemplo, dos *ratings*, uma vez que as formas de consumo mais intensivas e onde prevalece o “*zapping*” sobrestimam o seu valor. Também contribuem para a selecção dos suportes publicitários, permitindo identificar quais são os que desempenham o papel principal e secundário, bem como a intensidade das suas audiências conjuntas, isto é, partilhadas.

O princípio da regularidade da exposição sustenta as teorias clássicas e os modelos probabilísticos frequentistas de construção dos planos de *media* que recorrem aos inquéritos (*Zarate, 1997*). A maior variabilidade dos conteúdos e da oferta televisiva e a facilidade de selecção e transição entre os canais, poderá contribuir para uma maior irregularidade do consumo, inviabilizando determinadas perspectivas de análise quando se pretende realizar análises de previsão.

A selecção dos suportes também é função do perfil das suas audiências e dos comportamentos de “lealdade”, sendo esta questão abordada nas análises onde se pretende tentar associar conteúdos aos segmentos, e, assim, avaliar, também, se as influências estruturais permitem que eventuais preferências dos telespectadores se possam manifestar. Corresponde a uma etapa que antecede a construção dos modelos de exposição no decorrer do planeamento de uma campanha publicitária. A construção de um plano de *media* não corresponde a um processo totalmente automatizado e independente da intervenção humana. No decorrer deste processo o planificador tem que, frequentemente, realizar alterações, ajustamentos aos resultados que são produzidos pelos sistemas informáticos e fornecer-lhes determinada informação que será função do seu conhecimento a respeito da evolução dos dados e dos comportamentos de consumo televisivo.

Nomeadamente, quando se constrói um plano numa das etapas intermédias será necessário, atendendo ao elevado número de suportes, eliminar alguns e reter os que realmente permitem discriminar os vários segmentos alvo, facilitando uma planificação manual da campanha publicitária. Através da associação de suportes a segmentos contribuímos para uma melhor “filtragem” dos suportes, considerando as variáveis sociodemográficas que constam das bases de audimetria e avalia-se o perfil das audiências de determinadas tipologias e canais e a possível existência de “lealdade” aos suportes.

Para além de recorrer às variáveis de caracterização do painel que constam das análises realizadas, também, podemos incluir nas matrizes de exposição ou de índices de afinidade, a identificação dos públicos objectivo que são projectados nos planos factoriais conjuntamente com os suportes. Estas análises permitem associar segmentos e variáveis sociodemográficas aos suportes que vão constituir a amostra sobre a qual incidir as análises em etapas posteriores. Os métodos factoriais podem ser complementados com as Regras de associação, uma vez que os suportes devem no seu conjunto maximizar a exposição ao alvo e a capacidade de persuasão da mensagem a divulgar.

O público objectivo será definido com base nas variáveis sociodemográficas e psicográficas que podem ser especificadas em função das características do produto ou serviço a divulgar, mas também por vezes recorrendo a análises prévias de segmentação do painel e em função dos seus perfis de exposição ao *media*. Por vezes, é relevante comparar um plano construído manualmente com o que resulta da aplicação de determinados modelos matemáticos, antes de chegar à solução final e, neste contexto, as análises de duplicação contribuem para uma melhor selecção dos suportes televisivos. O plano final será função do orçamento disponível, da cobertura do segmento alvo e da intensidade da exposição, sendo, consequentemente, também, necessário nas etapas finais deste processo recorrer a algoritmos de optimização, onde o utilizador deverá definir as restrições existentes.

Em função dos resultados das análises realizadas, posteriormente, pretendeu-se verificar se é viável, a nível nacional, a construção dos planos de *media* recorrendo à previsão de valores de ratings programáticos e de ratings comerciais (*Questão 3*) e na avaliação das perspectivas de análise que em alternativa podem conduzir aos melhores resultados (*Questão 4*) e que são função dos comportamentos de consumo televisivo. A previsão dos ratings também é essencial quando se pretende adquirir suportes de forma individual no denominado “spot market”, onde as estações televisivas não garantem um valor mínimo de exposições ao pacote de anúncios contratado. A audimetria viabilizou a quantificação das audiências das inserções comerciais que não eram consideradas nos métodos tradicionais que recorrem aos inquéritos, assim como a duplicação das audiências.

Na presente tese, constam algumas análises na área da previsão dos principais indicadores de audimetria, recorrendo a análises univariadas e a análises de Regressão linear e não linear multivariadas de *Data Mining* ou de Análise Simbólica. As análises explicativas e de previsão dos indicadores comerciais também são relevantes, uma vez que os ratings comerciais correspondem ao indicador que realmente permite a comercialização das inserções, mas que depende fortemente da envolvente programática.

Em função dos resultados alcançados com a investigação empírica, no capítulo seguinte, consta um resumo das perspectivas de análise que podem viabilizar a construção dos planos e que são função dos contextos em análise.

Para além dos contextos que requerem ter em atenção a análise dos ratings e que foram realizadas na investigação empírica apresentada, também se descreve uma abordagem recorrendo aos dados dos ficheiros base de audimetria e ao *Data Mining* não proposicional que possa viabilizar uma construção mais eficiente dos planos, tendo em atenção os modelos explicativos da exposição que caracterizam em maior grau o painel de audimetria nacional. Entre a venda do espaço comercial e a transmissão da publicidade contratada decorre um curto espaço de tempo, o que levanta a necessidade das análises de previsão e de construção dos planos serem exequíveis em tempo útil. Por conseguinte, os modelos de *ratings* corresponde à perspectiva privilegiada pelo sector.

Em suma, o conhecimento antecipado dos valores dos indicadores de audimetria confere vantagens competitivas às estações televisivas, uma vez que a subestimação ou sobrestimação pode implicar a perda de receitas e, consequentemente, a rejeição do mercado. A previsão pode implicar e depender da caracterização e do desenvolvimento de modelos explicativos da exposição ao *media* e de selecção dos conteúdos. Em determinados contextos as previsões também podem ser viabilizadas através de uma mera análise da evolução histórica dos indicadores e da aplicação de metodologias de previsão de séries temporais. Os contextos condicionam as abordagens a adoptar, mas, independentemente, da via adoptada tradicionalmente tem-se considerado como essencial conseguir determinar a cobertura do mercado alvo e a frequência média de exposição ou os valores dos *ratings* que permitem o cálculo das audiências acumuladas de uma campanha publicitária.

V.2. ANÁLISE DA EVOLUÇÃO DOS INDICADORES AGREGADOS OU “GROSS MEASURES” DE AUDIMETRIA

Neste capítulo, pretende-se a avaliação da *Questão 1*:

Qual o poder explicativo dos factores que constam das bases de dados de audimetria ao nível das taxas médias de audiências para o total da TV, dos *ratings* programáticos e comerciais e dos “factores bloco” quantificados através da variável *Spot_Ratio%*?

Foram realizadas análises univariadas e multivariadas de caracterização da evolução dos principais indicadores de audimetria a nível programático e comercial e considerando como unidade de análise: uma faixa horária, a duração programática parcial e global, um espaço comercial e uma inserção publicitária. Segundo a reduzida bibliografia da especialidade, acessível aos investigadores académicos, as abordagens de previsão no contexto do planeamento de *media* incidem mais, frequentemente, na análise dos indicadores de audimetria, em vez de nas escolhas realizadas individualmente ou nos fluxos de audiências entre os suportes publicitários, uma vez que os comportamentos individuais são classificados como acausais e imprevisíveis.

Segundo *Taleb e Makridakis (2009)*, entre outros autores, a história não se repete da mesma forma e a extrapolação e a interpolação dos padrões históricos pressupõe que estes se mantêm constantes. Logo, será necessário avaliar o grau de estabilidade dos indicadores e as características das bases de dados de audimetria inviabiliza a aplicação das metodologias mais tradicionais de análise de dados.

Deve ser caracterizada a evolução dos principais indicadores e avaliar quais são os factores que os condicionam maioritariamente, devendo-se privilegiar a informação que consta das bases de dados de audimetria, uma vez que as análises devem ser realizadas com a maior rapidez e o menor custo possível e recorrendo aos dados que se encontram mais facilmente acessíveis. Conhecendo os factores que determinam a evolução das audiências programáticas poderá contribuir para que melhores resultados sejam alcançados quando se pretende estimar os seus valores futuros ou para uma construção das grelhas que maximize a exposição ao público pretendido, assim como os valores dos *ratings*. As variáveis a considerar podem corresponder a factores de planeamento ou de caracterização dos conteúdos, entre outras. Quanto maior for a dependência face aos factores estruturais de planeamento, mais facilmente, se poderá prever os indicadores

desde que exista alguma regularidade na exposição agregada ao *media*. Os modelos quantitativos de previsão das audiências são essenciais para o planeamento de *media*.

Podem surgir algumas dificuldades, quando se pretende utilizar algumas variáveis de caracterização da programação como variáveis de previsão, como, por exemplo, quando as análises incidem na identificação dos cenários históricos mais semelhantes à observação a prever. Por exemplo, quando as análises incidem no médio prazo e, por vezes, inclusive no curto prazo, poderá não ser conhecida a programação a transmitir dentro de algumas semanas. Porém, a nível nacional, uma vez que se transmitem essencialmente telenovelas e séries ou espaços informativos nos períodos de maiores audiências, esta limitação é atenuada, em grande parte do ano, uma vez definida a grelha no início da temporada televisiva.

Por vezes, quando existem alterações na grelha apesar dos conteúdos serem substituídos, mantêm-se as tipologias com o objectivo de fidelizar as audiências, como, por exemplo, nos espaços destinados a telenovelas e a “reality shows” ou concursos. Logo, foi necessário avaliar a variabilidade dos indicadores considerando diversos horizontes temporais que se diferenciam ao nível da constituição das grelhas e do desempenho global dos canais. Verificámos que em determinados períodos, as análises podem ser realizadas ignorando as características dos conteúdos, atendendo à evolução mais regular dos *ratings* e às influências estruturais na evolução dos dados.

O desconhecimento da grelha deverá corresponder em maior grau aos espaços de transmissão de filmes, uma vez que estes não são divulgados com a necessária antecedência, ou quando se transmitem determinados eventos, galas ou “notícias de última hora”. Por outro lado, a utilização das variáveis de caracterização dos conteúdos em contextos de maior variabilidade dos dados levanta inúmeras incógnitas do ponto de vista da sua capacidade predictiva. Quando se transmitem filmes, a exposição também poderá estar em menor grau condicionada por parte das características dos conteúdos, se existir uma lealdade tipológica ou motivações ritualísticas nas decisões de acesso ao *media*, o que permitirá reduzir a variabilidade dos indicadores.

Para alguns suportes publicitários poderá ser testada a análise dos cenários históricos mais comparáveis ao nível dos factores de planeamento, desde que imperem as motivações ritualísticas acompanhadas por alguma regularidade na exposição agregada, que tem que ser inicialmente avaliada.

Consequentemente, foi necessário avaliar a regularidade da exposição e recorreu-se a uma amostra que incide num horizonte temporal onde se verificam consideráveis alterações no desempenho dos canais, verificando como essas alterações se repercutem na variabilidade dos principais indicadores. Certamente, que sempre vão existir novos eventos que contribuem para uma maior aleatoriedade da evolução dos dados e o futuro é considerado como se diferenciando estruturalmente do passado, mas se a exposição estiver condicionada pelo factor hábito poderá ser viável realizar análises de previsão para alguns suportes, que são função das faixas horárias e tipologias em análise. Também, será necessário identificar os contextos de maior variabilidade dos valores e os factores que os condicionam.

Alguns autores defendem que a previsão dos índices de audiências poderá ser cada vez mais complexa (Weber, 2002), como consequência da fragmentação das audiências e da especialização que resulta de um acréscimo no número de canais, formatos de transmissão, novas formas de *media*, entre muitos outros factores.

- A fragmentação reflecte o declínio global dos valores dos *shares* como resultado do crescente número de alternativas e de uma exposição televisiva mais diversificada e menos concentrada nas diversas opções.
- A polarização das audiências e a especialização por contrapartida reflecte uma tendência para que as audiências de determinados canais e suportes serem cada vez mais exclusivas. Cada telespectador pode assistir a determinados conteúdos que o diferenciem significativamente dos conteúdos seleccionados pelos restantes telespectadores. As audiências são especializadas quando se verifica uma reduzida sobreposição entre as audiências de diferentes canais e programas.

A fragmentação das audiências tem sido observada nos USA (*Webster, 2005*) e em alguns países europeus (*Gerhards & Klinger, 2005*), como, por exemplo, na Alemanha e no Reino Unido. Contudo, as tipologias também divergem a este nível. Na Holanda, uma vez que o mercado informativo encontra-se menos polarizado do que nos USA, as audiências dos espaços informativos são menos fragmentadas e os telespectadores menos interessados em determinadas tipologias não mostram uma tendência tão significativa para evitar esses conteúdos quando cresce o número de alternativas disponíveis (*Wonneberger et al., 2012*). A exposição é maioritariamente linear e fortemente condicionada por específicas restrições estruturais (*Van der Burg, Lauf & Negenborn, 2011*). Em *Wonneberger et al. (2012)* conclui-se que as audiências mantêm o TV ligado como consequência de um entorno, um ambiente externo que o viabiliza e por uma maior relação para com o TV e o desenvolvimento de hábitos de consumo televisivo.

A nível nacional, será necessário avaliar se também se verifica esta tendência. O desenvolvimento de um modelo estrutural de caracterização da exposição televisiva consta da segunda questão em avaliação e refere-se às tendências observadas nos comportamentos de exposição televisiva do painel de audimetria nacional. Verificámos que, também, se observa uma tendência para alguma especialização das audiências, verificando-se alguma independência entre as audiências de cada canal para uma considerável fracção do painel. Quando cresce o número de alternativas, nem sempre as selecções apresentam uma significativa diversificação.

Porém, grande parte do painel nacional não tem acesso à TV Cabo e, consequentemente, apenas poderá assistir aos canais generalistas, acabando por se verificar alguma lealdade aos canais e a tipologias no decorrer do tempo. Também se observa que a exposição aos canais por TV Cabo nem sempre penaliza de forma muito significativa o desempenho dos canais generalistas. Consequentemente, será necessário avaliar a irregularidade que a contra-programação poderá incutir nos *ratings*.

Frequentemente, outras tecnologias de informação, como a Internet também têm sido apresentadas como complementares, não conseguindo retirar audiência ao meio televisivo (*Tewksburg, Weaver & Maddex, 2001*). Em princípio, deverá ser superior para grande parte do painel a “concorrência” exercida por outras actividades de lazer, em especial ao fim de semana ou em períodos de férias nas estações do ano menos propícias a actividades ao ar livre. Em muito maior grau, a exposição também estará condicionada pela possível “irregularidade dos tempos vazios” entre actividades laborais ou de lazer.

Será especialmente difícil realizar a previsão da exposição e dos indicadores que dela decorrem, independentemente das abordagens metodológicas que sejam utilizadas, se não existir o “hábito” de ver TV de forma regular em determinadas faixas horárias, e que foi avaliado através das análises de duplicação que foram realizadas, ou se os indicadores agregados também não apresentarem alguma estabilidade.

Verificámos que, frequentemente, os índices de duplicação entre os suportes publicitários são reduzidos, especialmente se considerarmos as análises de lealdade programática. Uma maior especialização das audiências poderá, eventualmente, contribuir positivamente para o grau de previsibilidade dos *ratings*. Porém, procurou-se relacionar a evolução dos indicadores agregados com os valores dos indicadores acumulados e verificamos que na realidade não existe uma relação proporcional entre eles, e, logo, por vezes, os *ratings* apresentam estabilidade apesar das reduzidas duplicações observadas nos comportamentos que os originam.

Uma vez que, também, se pretende promover a previsão de valores de *ratings* será necessário realizar análises exploratórias de caracterização da sua evolução, assim como de avaliação dos factores que condicionam, maioritariamente, a sua evolução e que constam das bases de audimetria. As análises exploratórias e explicativas, também, são necessárias para a definição da tipologia de modelos quantitativos a utilizar nas análises de previsão. (*Patelis et al., 2003*).

Igualmente no contexto de análise dos “*ratings*”, também, persistem inúmeras incógnitas quanto aos factores que condicionam e determinam a exposição televisiva, abordando-se alguns aspectos na *Questão 2*, assim como a sua associação e repercussão nos indicadores de audimetria (Webster, 2006). De acordo com Patelis et al. (2003), o consumo televisivo apresenta grande complexidade e os factores mais determinantes não estão bem definidos e hierarquizados.

No *Capítulo I* foi justificada a necessidade de recorrer a modelos de *ratings* que visam o comportamento agregado no âmbito do planeamento de *media*. Nas análises exploratórias e explicativas realizadas aos principais indicadores de audimetria pretendeu-se verificar se os programas ou as tipologias utilizadas pelo sector televisivo para classificar os conteúdos se diferenciam quanto ao grau de variabilidade que apresentam no decorrer do tempo ou perante alterações da sua posição na grelha de programas ou nos contextos da programação transmitida pelas estações televisivas concorrentes.

Poderá existir uma determinada tipologia de programas que em determinadas faixas horárias apresente uma maior variância dos principais indicadores de audimetria, colocando maiores obstáculos às análises de previsão e outros, por sua vez, podem ser mais estáveis.

Os contextos em análise e em função das suas características podem implicar, posteriormente, a necessidade de recorrer a diferentes abordagens. Por exemplo, por vezes poderá existir estabilidade nas taxas médias de audiências para o total da TV, sendo apenas necessário quantificar a sua repartição entre os diversos canais. Contudo, por vezes, nestes contextos também poderá existir “independência” entre os ratings de cada canal e nem sempre a contra programação poderá intervir como um factor determinante das selecções.

Consequentemente, são inúmeras as incógnitas que inviabilizam a apresentação de uma proposta metodológica que possa contribuir para a previsão, se inicialmente não forem realizadas análises de caracterização da evolução dos principais indicadores e da exposição televisiva. Neste sector, salienta-se, frequentemente, que os grandes inimigos da previsão, na actualidade, correspondem predominantemente às contínuas alterações que se verificam na grelha de programas e ao nível da contra programação, por vezes, realizada no dia anterior, assim como quanto aos horários de transmissão.

Foi necessário avaliar em que medida estes factores, entre muitos outros, contribuem ou não para a variabilidade observada nos principais indicadores de audimetria a nível nacional. No *Apêndice F.1* foram identificados os contextos de maior variabilidade de *ratings* e de taxas médias para o total da TV e verificámos que, nem sempre, as maiores alterações horárias implicam necessariamente uma maior variabilidade de valores. Na realidade as alterações horárias que em alguns dos anos considerados são frequentes nas principais faixas horárias, por exemplo, para o canal TVI e as alterações na contra programação nem sempre incutem nos dados de audimetria um significativo grau de variabilidade, podendo não transcender os 5%-10%, que alguns autores internacionais consideram como aceitáveis.

A variabilidade de alguns suportes, desde que excluindo observações atípicas, por vezes, não atinge os 5%, considerando um período característico do consumo médio. Frequentemente, a contra-programação não retira telespectadores de forma significativa a determinado canal, ao ponto de influenciar significativamente a evolução dos *ratings*, inclusive quando são transmitidos determinados eventos desportivos.

Os factores de disponibilidade e/ou “vontade” para ver TV ou um programa em concreto podem condicionar fortemente a evolução dos “*ratings*”. Consequentemente, a variável que, frequentemente, se destaca nas análises multivariadas lineares e não lineares realizadas de avaliação da associação existente entre os diversos indicadores de audimetria e os *ratings* corresponde à taxa média para o total da TV, que poderá, frequentemente, ser independente dos conteúdos que são transmitidos, uma vez que não

apresenta uma maior variabilidade quando se comparam os mesmos períodos para diferentes anos face aos valores de variabilidade em cada um dos períodos referente a cada ano considerado.

Foram realizadas análises explicativas da evolução das taxas médias de audiências para o total da TV e verificámos qual é a variabilidade deste indicador e quais são os principais factores, que constam das bases de dados de audimetria, que o condicionam. Sucintamente, verificámos que a evolução das taxas médias de audiências para o total da TV traduz a existência de um “padrão agregado”, mas, não, necessariamente individual de exposição televisiva que difere em função dos horizontes temporais diários e semanais considerados. Quando está em análise uma grelha para a qual já existe um histórico de valores verificamos que a programação poderá exercer alguma influência nos *ratings*. Porém, é muito mais evidente e considerável quando se verifica a transmissão de determinados eventos desportivos, essencialmente, alguns jogos de futebol quando envolvem os principais clubes nacionais ou a selecção. No entanto, meras tipologias programáticas e a informação que consta das bases de audimetria de caracterização dos conteúdos, por vezes, são insuficientes para se poder prever a sua repercussão neste indicador.

Nomeadamente, os jogos de futebol apresentam uma grande divergência quanto à extensão e dimensão da sua repercussão nos diversos indicadores, sendo a tipologia de programas que poderá colocar mais dificuldades quando se pretende realizar análises de previsão. No Reino Unido, foi realizada uma tese de doutoramento que apenas procura contribuir para a previsão das audiências dos jogos de futebol. A metodologia adoptada procura a identificação das características dos conteúdos que possam explicar a evolução dos indicadores e, assim, eventualmente, contribuir para a previsão das suas audiências (*Buraimo, 2007*).

Os eventos desportivos correspondem aos conteúdos onde será mais relevante a identificação de variáveis que os caracterizem, sendo esta menor para as restantes tipologias que constam das bases de audimetria, atendendo aos resultados das análises realizadas. Porém, segundo a bibliografia disponível, verificamos que, frequentemente, as variáveis identificadas não são previsíveis e dificilmente quantificáveis e imperam as contradições ao nível da sua influência na evolução dos dados. Logo, dificilmente as análises de previsão podem ser, eventualmente, viabilizadas através da identificação de variáveis de caracterização dos conteúdos.

Por exemplo, em *Reeth (2011)* constam inúmeras referências internacionais a respeito dos factores que condicionam a exposição à transmissão dos eventos desportivos, ficando bem patente a sua complexidade, a diversidade de factores e, por vezes, a falta de consenso quanto à existência e dimensão da sua repercussão. O autor salienta que a análise da repercussão da incerteza do resultado na exposição televisiva, apenas, mais recentemente, tem suscitado o interesse dos investigadores:

An important subject that sports economists analysing TV demand have concentrated their attention on is the impact of outcome uncertainty on TV viewership. Although there had been some previous research in this field, this line of research has only recently witnessed a significant boost in interest, kick-started by the seminal article of *Forrest, Simmons & Buraimo (2005)* on what they call the couch potato audience. (p.3, p.4)

Segundo este autor, *Forrest, Simmons e Buraimo (2005)*, concluem que a incerteza do resultado de um jogo condiciona moderadamente os seus resultados: “outcome uncertainty is a significant determinant of audience size, but the magnitude of its impact appears to be modest.” (*Reeth, 2011, p.4*). Entre os factores mais relevantes salienta a atractividade do jogo, a época, a incerteza do resultado do jogo, o canal (privados versus estatais), as equipas e a sua “qualidade” na medida em que conseguem proporcionar um bom espectáculo televisivo e os factores de planeamento. A atractividade de um jogo para cada um dos telespectadores, assim como o que cada um considera como um jogo de “qualidade” e a incerteza do resultado final que será função da evolução do jogo e, consequentemente, obviamente, desconhecido *a priori* dificilmente constituem variáveis de previsão, atendendo também à sua subjectividade.

Em análises de previsão mais facilmente se poderá incluir características mais objectivas, como, por exemplo, as equipas, a sua classificação nessa temporada e a fase de qualificação. Porém, estas variáveis apenas determinam uma probabilidade de selecção, a qual poderá não se verificar devido às inúmeras influências estruturais no consumo televisivo.

Nem sempre, a exposição pode ser associada a uma intencionalidade e a um interesse ou uma preferência. O consumo televisivo, também, se encontra, por exemplo, condicionado por parte das condições climatéricas, que não são facilmente previsíveis com a antecedência que é necessária ao sector televisivo e por factores pessoais ou do grupo em que o telespectador se encontra inserido. Se considerarmos outras tipologias programáticas, como, por exemplo, os filmes também se poderá considerar em análises de previsão a sua tipologia, o país de produção, a antiguidade, se foi previamente exibido no cinema, se existe em vídeo, quais são os actores principais, entre outras variáveis mais objectivas, mas, que na realidade também podem ser insuficientes.

A exposição face a um determinado suporte em concreto, também, não é o mais relevante num plano de *media*, mas, sim, o desempenho do conjunto de suportes. Ainda no campo desportivo, por exemplo, no estado da arte *Reeth (2011)* refere que:

Their work (*Forest et al. (2005)*) inspired similar studies on other European football leagues. For the Spanish first division *Garcia & Rodriguez (2006)* report that the attractiveness of a match is the main determinant of the size of the audience. They also find a strong seasonal component in the evolution of the size of the audience within the football season. Also *Buraimo & Simmons (2007)* analyse Spanish football and conclude that "unlike their stadium counterparts, who have a preference for outcomes which favour their home team, TV audiences overwhelmingly prefer close matches than ones in which the outcomes are more predictable. The Italian first division was studied by *Di Domizio (2010)*. He shows that more than 90% of the variability concerning TV audiences can be explained net of uncertainty factors. Variables associated with the closeness of the match increase the goodness of fit, but are not crucial in determining the TV share. Other European authors analysed TV football audiences from a different point of view.

In a novel approach, *Alavy, Gaskell, Leach & Szymanski (2010)* analyse minute to minute intragame viewership for English League football matches and find that "although uncertainty matters, it is the progression of the match which drives viewership and as a draw looks increasingly likely, viewers are likely to switch channels. Based on a comparison between two Scandinavian football leagues, *Johnsen & Solvoll (2007)* report that viewers have different

motives for watching games on a public service broadcaster to those watching on a private channel. For a private channel content (the types of games shown) is decisive, while for a public service broadcaster timing (scheduling strategies) is everything. *Nüesch & Franck (2009)* present evidence of patriotism in TV viewership by examining data for international football games between countries. They show that both the expected game quality based on the proven playing strength of the teams and levels of patriotism strongly predict the TV figures. (p.4)

Consequentemente, segundo inúmeros autores, as variáveis que são referidas como determinantes da exposição, nem sempre, são na realidade previsíveis e, consequentemente, outras perspectivas de análise devem ser consideradas quando se pretende a previsão da exposição televisiva e a selecção dos suportes a incluir nos planos. Nas Ciências Sociais, sendo uma ciência inexacta, impera a óptica explicativa da evolução dos dados e as previsões qualitativas e estocásticas. A análise dos factores mais qualitativos que podem determinar a exposição nestes contextos não foi abordada neste trabalho, por não ser o mais relevante quando está em causa a previsão das audiências numa perspectiva de construção dos planos de *media* e que se justifica na quarta questão em avaliação.

Em *Reeth (2011)*, refere-se, nomeadamente, que os factores de planeamento podem ser relevantes para os canais estatais e para os canais privados poderá assumir uma maior relevância determinadas variáveis de caracterização dos eventos desportivos. Porém, a nível nacional verificamos que se verifica o oposto, sendo os factores de planeamento essenciais para a análise e sustentabilidade dos fluxos de audiências entre os suportes publicitários se considerarmos os canais privados. Os canais estatais apresentam valores reduzidos dos principais indicadores e a exposição é mais selectiva e maximiza-se para ambos os espaços informativos, assim como para os programas adjacentes por efeito de inércia e de “lealdade” para com o canal, atendendo ao perfil sociodemográfico do seu público.

Também, se compararam os canais privados e os canais estatais ao nível da evolução da duplicação das audiências e da evolução das audiências comerciais e em todas as áreas foram identificadas algumas diferenças.

Foram realizadas análises com o intuito de verificar quais são as tipologias e os programas que incutem uma maior variabilidade no desempenho da programação dos canais concorrentes e verificamos que em geral apenas alguns programas desportivos apresentam nitidamente uma maior influência, assim como algumas galas ou eventos e, por exemplo, as transmissões associadas ao processo Casa Pia, entre outros processos judiciais, que suscitam um maior interesse na opinião pública, a Festa 1º Maio, entre outras festas populares, etc.

Aparentemente a variabilidade de *ratings* é, frequentemente, superior quando se altera a posição de um programa na grelha, do que perante alterações na programação concorrente se considerarmos uma grelha já existente em que a audiência já realizou as suas “escolhas” e se estabelecem fluxos concretos de audiências entre os suportes publicitários. A estruturação da grelha, do mercado e os canais são responsáveis por fluxos e comportamentos estruturais das suas audiências. A exposição não é apenas função de preferências, mas, também, resulta de inúmeros comportamentos e variáveis estruturais e dos hábitos desenvolvidos no decorrer do tempo, entre outras variáveis.

As alterações ao nível da contra programação que se realizam a muito curto prazo e que podem gerar observações atípicas, também, não são consideradas quando se recorre às sondagens de intenção de ver como método qualitativo de previsão das audiências, assim como quando se recorre a determinadas abordagens quantitativas.

Perante uma nova grelha de programas poderá existir uma redistribuição ao nível dos suportes mais seleccionados. Contudo, será função dos factores que condicionam maioritariamente a exposição de cada segmento de telespectadores. Os *ratings* em função das faixas horárias consideradas também podem não apresentar uma forte variabilidade. Os diversos conteúdos e suportes a que um telespectador se encontra exposto no decorrer do tempo condicionam-se reciprocamente, quer dizer, não são acontecimentos independentes. A rapidez com que se altera o conjunto de suportes também será variável de telespectador para telespectador e será função do horizonte temporal em análise.

Em suma, atendendo às inúmeras incógnitas que a evolução dos ratings apresenta é necessário realizar algumas análises exploratórias e descritivas da evolução dos indicadores que condicionam a venda do espaço comercial. As principais conclusões a reter descrevem-se nos seguintes capítulos.

V.2.1. Caracterização da evolução das taxas médias de audiências para o total da TV

De acordo com as análises realizadas, verificámos que esta variável apresenta alguma independência face ao conteúdo transmitido, desde que não se verifiquem consideráveis alterações na grelha de programas, nomeadamente a transmissão de programas com maior capacidade de adesão das audiências, destacando-se sem dúvida alguns jogos de futebol. A evolução desta variável, dia a dia, é consideravelmente semelhante, apresentando um primeiro período de maximização dos seus valores em torno ao espaço informativo da hora do almoço e um segundo período muito superior, que corresponde ao espaço informativo da noite e que se prolonga pelo horário nobre e ao início da noite. Porém, considerando os valores das taxas médias de audiências para o total da TV associados a programas em concreto, em vez de a períodos horários, verificamos que a variabilidade observada é muito significativa, podendo alcançar os 24% em alguns dos anos considerados.

A variabilidade é, essencialmente, influenciada pelas alterações verificadas nos horários de transmissão dos programas ou ao nível da contra programação quando se verifica a transmissão de programas que geram observações atípicas. A variabilidade pode ser muito significativa, inclusive em horizontes temporais que não estão influenciados pela sazonalidade que caracteriza a exposição televisiva. O consumo televisivo apresenta um forte carácter sazonal, contribuindo para uma ainda maior variabilidade dos dados de audimetria. Consequentemente, as análises foram restringidas a um horizonte temporal que é considerado como característico a nível médio do consumo televisivo pelo sector televisivo.

Com base nas análises de correlação das variáveis de audimetria verificámos que a variável *Tvr_Prog%* depende em maior grau dos factores temporais, podendo apresentar valores consideráveis de correlação quando se consideram as observações adjacentes. O período horário e se nos referimos a dias laborais ou ao fim de semana influência consideravelmente a variabilidade deste índice, que depende da disponibilidade e da vontade de ver TV do painel.

Considerando os segmentos horários, apenas à hora do almoço ou no período designado por horário nobre assistimos a uma maior estabilidade na evolução do índice quando consideramos cada um dos anos, correspondendo essencialmente à transmissão dos espaços informativos. Atendendo à evolução interanual verificamos que o diferencial de valores varia em função da distância entre os anos considerados, faixas horárias e se nos referimos ao fim de semana ou não. As faixas horárias não são equiparáveis quanto à sua propensão para se verificarem alterações nos padrões de disponibilidade ou de predisposição para ver TV do painel de audimetria, que, por vezes, poderá estar condicionada pelos conteúdos transmitidos e ser função da composição sociodemográfica do painel.

Apesar das alterações que se possam verificar na estruturação da grelha de programas, os períodos de maiores audiências da noite não são os que apresentam uma maior variabilidade de valores no decorrer dos anos. Os períodos mais penalizados correspondem à hora do almoço e ao decorrer da tarde em dias laborais, apesar das tardes serem tradicionalmente associadas a uma audiência de donas de casa e a um público que poderá apresentar uma maior disponibilidade. No decorrer da tarde, também podem estar presentes outros segmentos, como, por exemplo, quem exerce uma actividade em part-time, por turnos, profissões liberais, estudantes, desempregados, reformados, etc. que podem contribuir para esta maior variabilidade por existir uma exposição mais irregular e poderem usufruir de um maior tempo vazio que tende a ser irregular. As faixas horárias associadas a períodos que podem ser destinados a actividades de lazer e em que se poderá verificar uma maior disponibilidade da maior parte da audiência apresentam uma maior estabilidade de valores.

Apesar da grande proximidade nas evoluções médias deste indicador para os vários anos considerados, o diferencial de valores, de ano para ano, poderá ultrapassar os 10%. Portanto, basear as previsões deste indicador na média aritmética dos seus valores históricos para idênticos períodos horários não parece ser uma boa opção, sendo esta uma perspetiva de análise, frequentemente, utilizada na prática.

V.2.2. Caracterização da evolução das taxas médias de audiências ou *Ratings*

Considerando os valores dos *ratings* programáticos verificamos que o canal estatal RTP1 apresenta uma maior regularidade interanual e bimensal face aos canais privados generalistas, apesar de ser o canal onde no decorrer dos anos é menor a estabilidade nas características dos programas transmitidos. Aparentemente, existindo por parte do seu público-alvo uma selecção programática mais selectiva e que coincide com os dois espaços informativos do dia e os seus programas sucessores, este canal apresenta, consequentemente, uma certa regularidade ao nível dos *ratings* se excluirmos programas que geram observações atípicas. Em alguns dos anos considerados, os canais com uma posição mais consolidada diferenciam-se da TVI quanto aos programas e as faixas horárias que apresentam uma maior estabilidade de valores.

A evolução dos *ratings* em geral denota a possível existência de rotinas de consumo para, talvez, uma fracção considerável do painel, mas, que, posteriormente, não foi confirmada recorrendo às análises de duplicação das audiências. Não existe obrigatoriamente correspondência entre a evolução da exposição agregada e os comportamentos individuais, segundo se demonstra no *Apêndice E*. No decorrer da noite os programas vão apresentando valores de *ratings* de menor dimensão, mas, também, são menos variáveis, assim como se verifica no decorrer da tarde que apresenta uma maior estabilidade de valores. No horário nobre, os *ratings* são mais variáveis e, especialmente, para os programas transmitidos nestes períodos que apresentam uma menor exposição ao painel. O período horário associado a determinado programa é fundamental para se conseguir prever os intervalos de confiança dos valores de *ratings*.

Considerando os programas do horário nobre e da noite que maximizam os *ratings* para os diversos canais, verificamos que para todos os programas a variabilidade dos *ratings* é menor ao fim-de-semana. Quanto mais avançamos no decorrer da noite menor será, não só, a variabilidade como, também, os valores observados. O horário nobre corresponde a um dos períodos que apresenta uma maior irregularidade ao nível das taxas médias de audiências para o Total da TV.

Consequentemente, neste período, os *ratings* também são mais variáveis, especialmente se considerarmos os programas aos quais não se possa associar comportamentos de lealdade programática ou de outro cariz. Os principais factores de audimetria que condicionam a variabilidade de *ratings* são: o desempenho global do canal e a sua imagem, os dias sendo necessário, essencialmente, diferenciar a semana e o fim-de-semana, as faixas horárias, as alterações horárias e a programação desportiva ou alguns eventos.

Considerando os valores de *ratings* associados ao canal, ou seja, referentes a períodos horários e com base nas análises de clustering realizadas a *Rat_Período%* verificamos que os canais se diferenciam quanto ao grau de estabilidade e regularidade observada neste indicador. Por exemplo, o canal SIC apresenta uma nítida diferenciação entre as evoluções semanais e de fim-de-semana, contrastando com a RTP1 que apresenta evoluções mais semelhantes no decorrer de toda a semana.

A TVI apresenta uma maior irregularidade na evolução deste indicador no decorrer dos dias laborais ou de fim-de-semana, uma vez que em parte do período em análise ainda não apresenta uma posição muito consolidada no sector televisivo. Quando este canal se consegue começar a destacar assistimos a uma maior estabilização das suas audiências. A transmissão de eventos desportivos, quase sempre, tende a gerar evoluções mais atípicas da variável em estudo, estando a TVI inicialmente mais dependente da programação do canal RTP1, uma vez que os fluxos de audiências entre os dois canais privados generalistas já se encontram mais “consolidados” após os telespectadores realizarem as suas escolhas quando as grelhas e os conteúdos são alterados.

Conseguir prever como o painel vai reagir à introdução de determinado conteúdo depende dos factores que condicionam em maior grau a exposição, nomeadamente das variáveis e comportamentos estruturais e motivacionais ou de interesses ou preferências. Consequentemente, os valores de *Rat_Prog%*, por vezes, apresentam uma maior variabilidade em função das oscilações horárias e das durações programáticas, do que como consequência de consideráveis alterações nos comportamentos de consumo do painel para idênticos períodos horários no decorrer da semana. Se associarmos um valor de *ratings* a cada unidade temporal verificamos que, da mesma forma que sucedeu com a variável *Tvr_Período%*, existem períodos diferenciáveis quanto à dimensão e variabilidade e condicionados pelo canal a que correspondem.

No decorrer do dia, constatamos uma grande discrepância na evolução dos *ratings* e correspondente grau de variabilidade para determinados períodos horários. Aparentemente, existindo alguma estabilidade dos níveis de audiência para os períodos horários é viável pressupor alguma independência entre o público-alvo de cada canal face a uma grelha já conhecida, sendo a variabilidade de *ratings*, essencialmente, atribuída aos factores de disponibilidade e das decisões de acesso ao *media* e às alterações horárias e em menor grau aos fluxos de audiências entre os canais. Nestes contextos, é relevante uma correcta caracterização e quantificação da variável que representa a exposição global ao *media*.

V.2.3. A influência da programação na evolução dos principais indicadores de audimetria

A influência dos eventos desportivos na evolução dos principais indicadores de audimetria, apresenta uma grande variabilidade. Quando se transmitem determinados eventos, sejam estes desportivos, ou, não, a sua repercussão nos canais concorrentes poderá ser apenas ao nível de *shares* e nestes contextos podem-se manter os níveis dos *ratings* dos canais concorrentes ainda que os seus *shares* se reduzam.

Por vezes, os acréscimos nos valores de *ratings* não se devem à audiência que é retirada aos canais concorrentes, mas, sim, ao acréscimo que se verifica nas taxas médias de audiências para o total da TV. Isto poderá acontecer por apenas o público do canal exposto a esse programa ter conhecimento da existência dessa programação que é divulgada com a utilização de autopromoções e, como tal menos acessível ao público dos canais concorrentes.

Os telespectadores, em especial os que têm acesso à TV Cabo, podem apresentar um repertório de canais, nos quais incidem em maior grau as suas selecções facilitando, assim, o processo de selecção. Logo, a ausência de exposição poderá verificar-se como consequência dos comportamentos de lealdade, não sendo fácil alterar o conjunto de programas a que os telespectadores assistem com a maior regularidade que lhes é possível.

Verificámos que um programa terá que ser muito apelativo e diferenciar-se das características da programação que, em geral, é transmitida para conseguir retirar audiência aos restantes canais e será função da “qualidade” do jogo e da incerteza do seu resultado, entre outros factores. Os jogos de futebol menos relevantes podem conseguir atrair um maior número de telespectadores, que normalmente não assistem a esse canal nesses períodos horários, mas em menor grau a audiência dos canais concorrentes. A repercussão da transmissão dos eventos desportivos nos canais concorrentes, também, poderá ser apenas ao nível de *ratings*. Por vezes, os programas pela sua capacidade de adesão do público reduzem os *ratings* dos programas concorrentes e quase sempre correspondem aos jogos de futebol mais relevantes.

Nestes contextos, também, será necessário avaliar a amplitude do efeito, ou seja, se é apenas para os mesmos períodos ou se também se repercute nos programas adjacentes, como consequência dos efeitos de herança observados e que resultam da inércia do público, sendo superiores se o programa adjacente for um espaço informativo.

Os efeitos de herança ocorrem, essencialmente, se após o jogo for transmitido um espaço informativo, uma vez, que os telespectadores não tiveram possibilidade de assistir ao espaço informativo dos restantes canais por transmissão simultânea com o jogo. As motivações instrumentais associadas a esta tipologia de programas também pode incentivar à permanência do público que não costuma ver TV ou essas faixas horárias. Se o horário do jogo não coincidir com o horário dos espaços informativos em contra programação, a tendência será para parte da audiência assistir ao espaço informativo que lhe é mais familiar, habitual.

Segundo as análises realizadas aos indicadores acumulados, verificámos que existe lealdade ao canal. Consequentemente, a evolução dos *shares* de audiências, nem sempre, reflecte a evolução dos *ratings*, sendo, frequentemente, desnecessária a previsão do primeiro indicador quando se pretende quantificar o indicador que realmente condiciona a venda do espaço comercial.

A nível prático quando se pretende tentar prever a repercussão dos programas que geram observações atípicas, por vezes, consideram-se as situações históricas com as características mais semelhantes, e que são definidas em função do número de observações identificadas, e calculam-se determinados factores que denotem as correções a realizar nas evoluções médias. Consideramos que nestes contextos e quando se pretende a análise de diferentes cenários de contra programação é possível recorrer previamente a metodologias de *Data Mining*, por exemplo, às Regras de Classificação e de Associação em séries temporais multivariadas, identificando quais são as associações que penalizam ou não os canais concorrentes e qual a sua dimensão. Nos contextos em que programas de maior popularidade são transmitidos, como, por exemplo, alguns jogos de futebol, devem ser realizadas análises mais exaustivas, recorrendo a metodologias da área do *Data Mining* Temporal. Esta questão desenvolve-se do ponto de vista metodológico no *Capítulo VI*, mas as previsões apenas serão viáveis se existir uma maior influência das variáveis estruturais e se existir alguma regularidade na evolução destes indicadores.

V.2.4. A evolução dos *shares* de audiências

A variabilidade ao nível dos *shares* é consideravelmente mais significativa, uma vez que esta variável encontra-se, fortemente, influenciada por variações nas taxas médias de audiências para o total da TV. Uma maior estabilidade nos valores dos *ratings* poderá não ser acompanhada ao nível da evolução dos respectivos *shares*. Portanto, não consideramos que em inúmeros contextos os *shares* constituam uma boa variável a ser incluída em análises de previsão, assim como, *Danaher et al.*, (2011), também, o referem num trabalho de previsão de *ratings* realizado com bases de dados australianas.

V.2.5. A segmentação da emissão diária

Realizando algumas análises de segmentação, canal a canal, é viável avaliar a extensão de cada um dos períodos em que os valores dos *ratings* apresentam uma maior estabilidade e equivalente grau de variabilidade e, assim, retirar ilações a respeito do possível grau de duplicação de audiências. Estes resultados foram confirmados no capítulo dedicado à análise dos efeitos de herança entre programas e faixas horárias (*Apêndice F.4*) e aos fluxos de audiências entre os suportes. Por conseguinte, em primeiro lugar analisaram-se os *ratings*, canal a canal e, posteriormente, simultaneamente por forma a identificar quais são as faixas horárias mais comparáveis ao nível da evolução do indicador que condiciona maioritariamente a actividade de planeamento de *media*.

Verificámos que o canal estatal RTP1 apresenta, frequentemente, como período de maiores audiências um horizonte temporal mais alargado e que tem início com o espaço informativo da noite, incluindo, também, o programa sucessor. Os canais privados generalistas podem apresentar maiores índices para as respectivas telenovelas da noite, seguindo-se os espaços informativos, por vezes, com índices ligeiramente inferiores. Por aplicação de algoritmos de *clustering* simbólicos e para conseguir percentagens de variância explicada equiparáveis, os canais privados generalistas superam o canal estatal RTP1 em número de agrupamentos. Os canais privados apresentam um maior número de segmentos temporais com evoluções diferenciadas.

Posteriormente, segmentaram-se os períodos horários com base na totalidade dos canais por aplicação de Redes Neurais e algoritmos de *clustering* simbólicos. Para as Redes Neurais simbólicas resultam quatro agrupamentos como a melhor solução e que abrangem períodos horários não contínuos. Em relação aos valores de variabilidade, estes são muito consideráveis, considerando esta partição da emissão diária.

A previsão de valores de *ratings* deverá ser, essencialmente, realizada em função dos períodos horários e considerando reduzidas unidades temporais, nomeadamente equivalentes ou inferiores aos 30 minutos e não se deve recorrer a programas em concreto, especialmente para os dias laborais. Posteriormente, com base nestes valores podem ser calculados os valores programáticos em função dos períodos horários que estes abrangem.

V.2.6. A INFLUÊNCIA DAS VARIÁVEIS SOCIODEMOGRÁFICAS NA EVOLUÇÃO DOS PRINCIPAIS INDICADORES DE AUDIMETRIA

Após as análises de segmentação, verificámos que os índices de audiências dos segmentos sociodemográficos não permitem discriminar os diferentes períodos. Os segmentos considerados nas bases de dados de audimetria apresentam valores muito correlacionados. Por aplicação da análise ACP simbólica apenas é possível retirar um factor, logo, os vários segmentos apresentam um comportamento semelhante do ponto de vista agregado.

Na avaliação da influência das variáveis sociodemográficas na evolução de *Tvr%*, verificamos que os maiores valores e, superiores ao universo televisivo, podem ser associados a: região do interior, classe social D, sexo feminino, faixas etárias superiores aos 55 anos e os lares sem TV Cabo. Com menores valores podemos destacar a classe social AB e os Lares com TV Cabo. Em dias laborais, existe uma forte semelhança entre a ordenação relativa dos diversos segmentos em função do seu grau de exposição ao *media*, sendo a classe social e as faixas etárias as variáveis que os permitem diferenciar em maior grau.

Em relação às análises sociodemográficas foi necessário verificar para cada um dos canais quais são os segmentos que apresentam evoluções dos *ratings* mais próximos entre si ou face ao universo televisivo. Em futuras análises de previsão será necessário ter em atenção a composição sociodemográfica das faixas horárias e em especial ao nível dos segmentos que apresentam padrões de consumo mais variáveis para cada um dos canais.

Os segmentos com padrões de exposição mais irregulares justificam a sua posterior análise recorrendo aos dados dos ficheiros de base de audimetria, uma vez que contribuem para uma maior variabilidade dos principais indicadores. Os segmentos com padrões de consumo mais atípicos devem ser analisados recorrendo a bases de dados transaccionais e a algoritmos de *Data Mining* Sequencial ou de Regras de Dependência e Associação, por forma a caracterizar quais são as selecções programáticas mais prováveis em determinados contextos.

Apenas nas faixas horárias que, realmente, apresentem valores mais variáveis dos *ratings* devemos considerar a análise dos dados de audimetria para segmentos ou a nível individual, tentando perceber a dinâmica das escolhas realizadas e a repercussão dos inúmeros factores estruturais, entre outros. Nestes contextos, em primeiro lugar devem ser previstos os valores de ratings em função da sua associação com as restantes variáveis de audimetria, nomeadamente com o indicador que representa as taxas médias de audiências para o total da TV. No *Apêndice F.1* também consta a identificação dos macro-segmentos em função da evolução dos *ratings* por segmento horário e canal considerado.

V.2.7. A análise aos *ratings* programáticos: Os modelos explicativos da evolução dos *ratings* – A extracção de regularidades

Para as análises multivariadas dos indicadores que constam das bases de audimetria e atendendo à globalidade dos dados a variável explicativa mais linearmente correlacionada com os *ratings* programáticos corresponde às taxas médias de audiências

para o total da TV, sendo este valor ainda superior se considerarmos os valores associados a períodos horários e em especial para os canais privados generalistas. Também, corresponde ao indicador que se associa em maior grau de forma não linear com os valores dos *ratings*. O canal estatal, RTP1, uma vez que apresenta uma programação mais diversificada, transmitindo inúmeros eventos desportivos, gera observações atípicas, com elevados valores de *ratings*, reduzindo a linearidade existente entre as variáveis mencionadas.

Algumas variáveis temporais, como, por exemplo, as horas de início ou de conclusão de um suporte publicitário não apresentam coeficientes de correlação linear que possam ser considerados relevantes. Uma vez que verificamos a existência de uma forte associação linear e não linear entre os valores dos *ratings* e as taxas médias de audiências para o total da televisão, frequentemente, será desnecessária a prévia previsão dos valores de *shares* de audiências que, por vezes, apresentam uma maior variabilidade de valores. Logo, a análise da repercussão da contra-programação nos *shares* de audiências, nem sempre, será previamente necessária.

A previsão das taxas médias de audiências para o total da TV poderá viabilizar a posterior previsão dos *ratings*. Consequentemente, também, foram realizadas algumas análises exploratórias a este nível, mas os resultados nem sempre são satisfatórios, especialmente se considerarmos o fim-de-semana. Obviamente, que os factores de disponibilidade ou de predisposição para ver TV apresentam uma inegável repercussão nos resultados cuja variabilidade poderá ser atenuada quanto mais o consumo se basear em rotinas desenvolvidas ao longo do tempo.

Porém, atendendo às análises realizadas no âmbito da lealdade programática, uma menor percentagem do painel e que será função das faixas horárias em análise poderá apresentar uma exposição televisiva associada a algumas rotinas de consumo televisivo.

Apesar da associação linear existente entre as variáveis *Tvr_Prog%* e *Rat_Prog%*, apenas em alguns períodos se mantém esta relação, em função do tempo, considerando como variável explicativa a hora de início do programa. Verifica-se para os períodos da manhã, início da tarde e final de tarde. A partir do pré prime time altera-se esta relação, uma vez que é maior a dimensão da audiência e com uma composição sociodemográfica semelhante à do universo televisivo. Verificámos que os resultados das análises de Regressão estão, essencialmente, associados à evolução semanal dos dados, com declives acentuados nos períodos adjacentes aos espaços informativos. Ao fim-de-semana, a evolução é menos irregular, com o período da manhã a apresentar uma certa estabilidade de valores, assim como a tarde.

Os valores de *shares* não estão linearmente correlacionados com os respectivos valores de *ratings* para nenhum dos canais considerados, atingindo no máximo um valor de $\rho=0.5$.

Existe uma tendência no público para assistir à televisão em sessões, até estarem satisfeitas determinadas necessidades, sendo estas de preenchimento de tempos livres, tempo vazio, entre outras. Consequentemente, antes e após assistirem ao seu programa mais habitual, em função da influência das variáveis estruturais, permanecerão em frente ao televisor por um determinado período, variável em função do telespectador. As observações adjacentes podem apresentar alguma correlação, contudo existe uma maior associação entre o *rating* de um programa e o *rating* do seu programa sucessor, também designado de “Lead_Out” face ao programa que lhe antecede.

Se considerarmos os valores referentes aos períodos horários, os valores de correlação tendem a equiparar-se. A variável *Rat_Ldout%* é mais relevante, ou seja, existe um fluxo de audiências a partir de um programa para os programas sucessores. Isto permite-nos reforçar a ideia de “passividade” da audiência, assistindo-se a um programa simplesmente porque sucede outro de maior interesse ou mais habitual. Os programas que podem apresentar maiores associações com os seus valores de “Lead_In”

correspondem, essencialmente, aos que sucedem os espaços informativos ou os programas associados ao período que decorre desde o horário nobre até às 22H-23H.

O canal RTP1 corresponde ao que apresenta maiores valores de fidelidade e de *ratings* para os espaços informativos, mas, também, é o que menos favorece os programas adjacentes, reduzindo a correlação existente entre os suportes adjacentes. No período da noite, encontra-se favorecido o programa que antecede o espaço informativo com maiores níveis de exposição para os vários segmentos, assim como o programa sucessor. A notoriedade do canal influencia o desempenho dos programas, sendo para o período em análise, os ratings superiores para os canais privados, com excepção dos eventos desportivos do canal RTP1.

As sinergias entre as tipologias dos programas não condicionam os resultados, atendendo às análises realizadas. Por vezes, é salientado que a preferência e a exposição a determinadas tipologias poderia favorecer os índices de audiência, desde que o planeamento favorecesse a transmissão contínua das tipologias mais “apelativas” à maioria do público, e, assim, incentivar a maiores índices de duplicação de audiências entre os programas adjacentes.

No decorrer do horário nobre transmitem-se em primeiro lugar os programas informativos, seguindo-se os programas de ficção para os canais privados, com a transmissão de séries, telenovelas, humor, concursos e filmes. A RTP1 apresenta uma planificação da grelha mais versátil, transmitindo programas de humor, concursos, programas desportivos, filmes, programas de debate, entre outros. Os programas que não consistem em retransmissões e que são transmitidos em séries apresentam os maiores valores de exposição televisiva.

Com base nestes resultados, verificámos que existem limitações quanto ao desenvolvimento de modelos de Regressão linear múltipla, considerando a globalidade dos dados.

Um reduzido número de variáveis apresenta alguma correlação com a variável em estudo, a maioria das quais não é conhecida no momento de realizar uma previsão. Todavia, se restringirmos as análises a períodos mais restritos, talvez, se possam alcançar resultados mais satisfatórios. O coeficiente de determinação apresenta o valor de 0,82 para o canal SIC, considerando apenas *Rat_Prog%* e *Tvr_Prog%* e ascende a 0,80 para a TVI e a 0,83 para a RTP1. Considerando a globalidade dos dados, não identificamos mais nenhuma variável de audimetria que deva integrar um modelo explicativo ou de previsão linear.

Restringindo as análises a períodos mais restritos, destaca-se o valor do *share* e as taxas médias de audiências para o total da TV. As restantes variáveis não viabilizam a construção de modelos com consideráveis valores dos coeficientes de determinação. Verificámos que os modelos de Regressão linear múltipla simbólicos identificam os factores temporais e em especial a variável *Tvr_Prog%*, como sendo as variáveis que mais se associam aos *ratings*. A variável nominal que representa o período horário apresenta um menor coeficiente de correlação. Logo, a previsão das taxas médias de audiências para o total da TV é fundamental para conseguir prever os valores dos *ratings* programáticos. Os canais privados apresentam evoluções de ratings mais semelhantes entre si e podem ser fortemente penalizados com a transmissão de eventos desportivos por parte do canal estatal RTP1. Excluindo os programas de grande popularidade, a TVI corresponde ao canal que poderá penalizar em maior grau o desempenho do canal SIC.

Realizando uma análise de Regressão linear Simbólica á variável dependente *Rat_Prog%* e considerando como variáveis explicativas *Tvr_Prog%*, o *Período* e *Tempo_Time*, verificamos que a variável *Tvr_Prog%* é a que apresenta um coeficiente significativo e ao modelo corresponde-lhe um $R^2=0.87$. Apesar dos coeficientes de determinação poderem ser superiores para as abordagens simbólicas, as variáveis com maior capacidade explicativa, também, correspondem aos factores temporais.

As análises de Regressão não linear apresentam uma capacidade explicativa entre os 80% e os 87% para a totalidade dos dados. Da aplicação das Árvore de Regressão CART e CHAID aos valores dos *ratings* programáticos interessa reter que as audiências também são essencialmente função dos factores temporais e das taxas médias de audiências para o total da TV. Os intervalos de *Tvr_Prog%* diferenciam-se para ambas as metodologias, no entanto, estas coincidem em que estes valores condicionam as variáveis que mais associam aos valores dos ratings programáticos. Para os períodos de maiores audiências assumem uma maior relevância os valores dos *ratings* adjacentes, alternando-se os anteriores com os sucessores. Em períodos de tendência crescente os *ratings* estão mais associados aos valores sucessores. Em períodos de maior estabilidade ou tendência decrescente assume um maior destaque os valores anteriores.

Em períodos de menores audiências e para os canais estatais deverá existir uma selecção programática mais pontual, e assim sendo, os *ratings* estão em maior grau associados a tipologias e, não, aos programas adjacentes por inexistência de relação entre os seus *ratings*. Os factores temporais influenciam os resultados e em virtude da evolução gradual das audiências, também, é notória a repercussão dos valores adjacentes que em função dos períodos horários se alternam em relevância e quanto a uma maior similitude de valores com a envolvente. Verificámos que os índices de duplicação entre programas directamente adjacentes podem ser razoáveis em virtude da existência de proximidade entre os valores. Obviamente, que em períodos de tendências ascendentes a incorporação gradual de novos elementos do painel será o factor responsável por tal evolução, uma vez que os índices de duplicação são sempre inferiores à unidade.

A capacidade explicativa das Árvore de Regressão, atendendo aos dados dos canais privados generalistas, apresenta valores entre os 80% e os 90% e para os canais estatais os valores aproximam-se dos 80%. O melhor modelo identificado para os canais estatais pode ser utilizado em análises de previsão desde que seja viável prever o valor de *Tvr_Prog%*, mas a capacidade explicativa também é de apenas 80%. No contexto dos canais privados para alcançar os 90% seria necessário prever *Tvr_Prog%* e a fidelidade ao programa.

A capacidade de previsão, consequentemente, fica consideravelmente aquém da explicativa. Nestas Árvores de Regressão, as previsões coincidem com a média das observações inseridas em cada nó final da Árvore. Apesar da frequente elevada capacidade explicativa esta não parece ser a melhor opção pelos factores já enumerados, assim como atendendo a que os intervalos de *Tvr_Prog%* que, por vezes, surgem no desenvolvimento das Árvores, têm amplitudes consideráveis e, por conseguinte abrangendo vários períodos horários, nem sempre adjacentes e sequenciais.

Os períodos horários influenciam as audiências e apresentam diferentes evoluções consoante a sua localização na emissão diária. Por exemplo, nas análises lineares foi possível constatar a existência de relações localmente lineares entre os *ratings* e *Tvr_Prog%* e as observações apresentam uma maior semelhança quanto mais próximas estiverem.

As Árvores de Regressão Local ou as Árvores cujas previsões são calculadas com base em metodologias de KNN também parecem ser uma opção a considerar. Alternativamente, teriam que ser aplicadas as metodologias a cada subconjunto de dados incluído em cada um dos nós da Árvore. Verificámos, também, que as variáveis comerciais não apresentam relevância quando se analisa a evolução dos *ratings*, inclusive quando reduzimos as análises a períodos horários mais restritos, assim como as variáveis de caracterização dos programas, entre as quais as seguintes: a duração de um programa ou dos programas adjacentes, os índices de saturação publicitária, se os programas consistem em retransmissões ou se os programas são transmitidos em episódios, as sinergias entre tipologias de programas, entre muitas outras variáveis, que foram calculadas para enriquecer as bases de audimetria e os resultados das análises.

V.2.8. A análise dos indicadores comerciais: Os modelos explicativos da evolução dos *ratings* comerciais a nível absoluto e relativo: A extracção de regularidades

Devido à relevância prática da caracterização da evolução dos *ratings* comerciais e

correspondente previsão, realizaram-se algumas análises aos índices *Rat_Spot%* e *Spot_Ratio%*, ou seja, aos índices que condicionam a venda das inserções publicitárias, assim como aos seus valores quando são definidos proporcionalmente face às correspondentes audiências programáticas. Para além das análises ANOVA realizadas e que permitiram avaliar a repercussão das variáveis nominais na evolução destes dois indicadores também se avaliou a linearidade das relações existentes entre os índices comerciais e as variáveis estruturais que constam das bases de audimetria por aplicação da Regressão linear múltipla e da Regressão linear múltipla simbólica. As análises não Lineares de Regressão resultam da aplicação das Árvores de Regressão. Pretendeu-se verificar se os factores de planeamento condicionam a evolução das audiências comerciais, ou se os *ratings* são apenas o reflexo das audiências programáticas que lhe são adjacentes e em que medida as alterações na grelha podem incutir-lhe variabilidade.

Os *ratings* comerciais devem ser uma função dos valores dos *ratings* programáticos, mas a associação deverá ser diferente em função das faixas horárias consideradas, dias e canais. Segundo as análises exploratórias realizadas, apesar do canal estatal RTP1 apresentar um perfil de audiências que em princípio apresenta uma menor predisposição para o “zapping”, o decréscimo nos valores dos *ratings* no decorrer de um espaço comercial a nível médio é significativo, podendo atingir um valor de 70% face ao valor das audiências programáticas. Apenas as duas primeiras posições e a última apresentam menores penalizações, atingindo a última inserção um valor de 86% do valor programático que lhe está associado em média.

Os espaços comerciais localizados entre programas perdem menos audiência e apenas as duas posições iniciais são favorecidas, enquanto, os intervalos entre programas apresentam maiores valores para estas posições e também para a última. Os programas de maior popularidade, por exemplo, os eventos desportivos conseguem manter um fluxo de audiências mais uniforme no decorrer dos seus espaços comerciais. As inserções transmitidas no decorrer da tarde e do horário nobre são as que apresentam uma menor penalização, seguem-se as inserções da noite e por último do pré prime time e da madrugada.

Assim como, se verificou para as análises de índices de lealdade programática, em horários mais tardios não existe obrigatoriamente uma audiência de menor dimensão, mas mais regular e mais passiva no decorrer dos espaços comerciais. O horário nobre é favorecido pela exposição da audiência no início e no final do espaço comercial.

Para o canal RTP1, apesar do decréscimo das audiências comerciais, verificamos que existe proporcionalidade entre a evolução dos *ratings* programáticos e os valores comerciais. A nível médio não podemos afirmar que exista uma menor redução dos *ratings* comerciais durante o fim-de-semana face aos dias laborais. Interessa, sobretudo, realçar que a redução ao nível dos índices de audiências das posições intermédias é superior ao fim de semana.

Considerando as tipologias programáticas, as categorias que apresentam menores penalizações correspondem aos documentários, programas de política, reportagens, programas de variedades, filmes western e os *sketches* humorísticos. Perante o horizonte temporal em análise, as promoções de programas não influenciam os resultados. Idênticas análises foram realizadas para os ratings comerciais e para cada um dos canais privados generalistas. Segundo *Toro (1997)*, que realizou, em Espanha, uma tese de doutoramento que visa a análise aos comportamentos de “zapping”, não existe um conjunto identificável de variáveis de audimetria que permitam explicar os decréscimos observados nas audiências comerciais, sendo este um factor essencial para o planeamento de *media* e, conseqüentemente, um tema de investigação que deve ser aprofundado. Este autor refere as seguintes considerações:

Después de analizar el nivel de fugas de audiencia producidas por cada una de esas distintas variables, podemos afirmar que el comportamiento de la audiencia de televisión depende de un conjunto muy amplio y diverso de factores, de tal magnitud, que ha sido imposible, incluso con la muestra utilizada por nosotros, llegar a determinar cómo varía dicho comportamiento en función de cada una de las variables estudiadas una a una; incluso analizando todas y cada una de las posibles interrelaciones entre todas las variables estudiadas, nos ha sido imposible determinar la existencia de una combinación que explicase de forma directa el comportamiento de la audiencia. Es decir, ninguna variable por sí sola, ni ninguna de las formas posibles de combinación de ese conjunto de variables, ha sido capaz de establecer el comportamiento general de la audiencia.

Este resultado puede deberse a dos razones principales: que existe una variable o una serie de variables no estudiadas (por no estar recogidas en la muestra audimétrica) que influyen determinantemente en el comportamiento de la audiencia, o que este comportamiento es mucho más rico y complejo que el que puede llegar a determinarse por la interrelación de una serie limitada de variables, por muy extensa que ésta sea. En cualquier caso, y desde el punto de vista de la realidad analizada, el fenómeno del cambio de canal ante los bloques publicitarios por parte de un porcentaje nada despreciable de la audiencia, debe estar muy presente en la planificación de medios del anunciante, y es un tema lo suficientemente importante como para seguir invirtiendo recursos en la investigación iniciada con este trabajo. (p.1)

Com dados nacionais, e com base nos resultados apresentados no *Apêndice F.7* verificámos quais são as variáveis que condicionam a evolução dos *ratings* comerciais a nível absoluto ou relativo. Diferenciaram-se os resultados para os vários canais generalistas, sendo as variáveis mais relevantes determinados factores de planeamento, como, por exemplo, a posição da inserção publicitária, a duração acumulada desde o início do intervalo e do programa e a localização do espaço comercial, assim como a faixa horária considerada e as tipologias programáticas envolvidas. Observaram-se consideráveis interacções entre as variáveis de audimetria que apresentam uma considerável capacidade explicativa da evolução dos *ratings* comerciais, mas muito menor para *Spot_Ratio%*. Verificamos a existência de significativas interacções entre as variáveis explicativas e que se encontram dependentes do canal em análise.

Var._Explicada	Rat_Spot (%)	Spot_Ratio (%)
Canais Públicos		
Global	0,82	0,25
Prime Time + Noite	0,76	0,48
Pre Prime Time	0,52	0,35
Tarde	0,76	0,53
Canais Privados		
Global	0,87	0,22
Prime Time + Noite	0,85	0,43
Pré Prime Time	0,54	0,31
Tarde	0,8	0,52

Para os canais estatais, os modelos de Regressão linear múltipla apresentam coeficientes de determinação consideravelmente superiores para os *ratings* comerciais face aos que foram calculados para o indicador *Spot_Ratio%*, atingindo o valor de 0,82 para a totalidade da amostra e de 0,76 para a subamostra que apenas incide nos espaços comerciais após as 20H ou no período da tarde.

Tabela V.2.9.1 - Coeficientes de determinação dos modelos de Regressão Linear.
Fonte: Elaborado pela autora.

O pré prime time apresenta o menor valor, sendo este de apenas 0,52. Para ambos os canais privados generalistas mantém-se a hierarquização dos períodos horários em função da ordem de grandeza dos coeficientes de determinação. Os modelos de Regressão linear múltipla para a amostra global atingem no máximo um valor de 87% de variância explicada e quando consideramos as inserções após as 20H este valor reduz-se ligeiramente para 85%. Considerando os valores dos *ratings* comerciais as análises de RLM e em especial as análises de Regressão linear simbólicas constituem uma alternativa às análises de Regressão não linear com Árvores de Regressão, apresentando ambas as metodologias razoáveis percentagens de variância explicada para os diversos horizontes temporais em análise.

Um modelo de Regressão linear Simbólico para a variável *Rat_Spot%*, incluindo apenas as variáveis que podem ser conhecidas *a priori* ou estimadas ainda que com alguma margem de erro que se pretende que seja a menor possível, por exemplo, *Tvr_Prog%* que apresenta uma menor variabilidade nos seus valores tem um coeficiente de determinação de 0,87. Porém, em alguns contextos será necessário avaliar a capacidade de previsão destes modelos que poderá ficar, consideravelmente aquém destes valores, uma vez que, por vezes, incluem variáveis que não são conhecidas.

As variáveis de audimetria apresentam alguma capacidade explicativa da sua evolução se considerarmos os valores absolutos das audiências comerciais, mas em muito menor para *Spot_Ratio%*. Os modelos de Regressão linear, neste último contexto, apresentam no máximo para a amostra global dos canais estatais um coeficiente de determinação no valor de 0,25 e de 0,22 para os canais privados. Se restringirmos as análises às faixas horárias da noite dos canais estatais este valor não transcende os 0,48 e para os canais privados é de 0,43. Verifica-se uma forte associação linear e não linear entre os *ratings* programáticos e os *ratings* comerciais. Contudo, foram alcançados resultados divergentes para o indicador *Spot_Ratio%*, para o qual as análises de Regressão linear e não linear não permitiram atingir resultados satisfatórios, que possam viabilizar a previsão deste indicador ou a identificação das variáveis que condicionam a sua evolução.

Neste contexto, subscrevemos os argumentos apresentados em *Toro (1997)*. Para a variável *Spot_Ratio%* se considerarmos apenas as variáveis conhecidas ou que podem ser calculadas *a priori* obtemos um R^2 no valor de 0,39, ou seja, muito menor ao do modelo associado a *Rat_Spot%*, mas superior ao dos modelos lineares clássicos.

Relativamente às análises não lineares verificamos que para o canal RTP1 as percentagens de variância explicada oscilam entre os 94% e os 80%, se considerarmos os ratings comerciais e para *Spot_Ratio%* ao melhor modelo corresponde 52%. Para os canais privados generalistas a SIC apresenta o valor de 90% e de 50% respectivamente e a TVI o valor de 87% e 48%. Os índices comerciais são função dos índices dos programas adjacentes, mas como ficou demonstrado não corresponde simplesmente ao preenchimento da audiência entre dois programas. A evolução dos *ratings* no decorrer dos espaços comerciais poderá ser linear ou não linear em função dos factores de planeamento.

Segundo as análises lineares, para o canal SIC algumas das variáveis que penalizam os ratings comerciais são a saturação publicitária de um programa e em maior grau do seu antecessor; a duração de um programa global ou do programa anterior e inclusive se considerarmos o número de intervalos que os integram; a duração desde o início de um programa até à parte associada a determinado espaço comercial a nível absoluto e relativo e número de promoções em um intervalo. Com uma associação positiva destacamos os valores dos ratings programáticos e dos programas adjacentes ou os níveis das taxas médias de audiências para o total da TV, assim como a fidelidade programática.

A identificação de quais são os factores que se associam e condicionam *Spot_Ratio%* é de grande relevância e devem-se calcular estes factores para cada canal, fracção do dia e público objectivo. Segundo as análises lineares, para o canal SIC apesar do reduzido valor destes coeficientes verificamos a existência de uma associação positiva com a fidelidade programática e com *Average Time Viewing (ATV)* e a hora de transmissão da inserção.

Com valores negativos destaca-se a duração de um programa ou do número de intervalos que o integra, o número de promoções, a duração desde o início do programa e a parte que lhe está associada, a duração desde o início do intervalo e a transmissão da inserção a nível absoluto ou relativo face à duração global do espaço comercial, os ratings programáticos e os níveis das taxas médias de audiências para o total da TV, o número de inserções e a duração do espaço, a posição da inserção e a duração da inserção, entre outras variáveis de planeamento.

Verificamos que a variabilidade dos *ratings* comerciais nem sempre é elevada, por exemplo, para o canal SIC em alguns dos anos considerados não transcende o valor de 7%, sendo este valor de 6% para a RTP1 se excluirmos determinados eventos desportivos. Uma vez previstas as audiências programáticas podem ser calculadas as audiências comerciais, atendendo à significativa relação linear e não linear observada entre ambas as variáveis para a maioria dos contextos em análise. Por exemplo, considerando a amostra global do canal SIC, o coeficiente de correlação é de 0,79 e o modelo alternativo mais parcimonioso apresenta um coeficiente de determinação no valor de 0,80, incluindo também a duração do programa e o número de promoções transmitidas no decorrer do espaço comercial.

Do ponto de vista comercial não se deve realizar a previsão para blocos de anúncios, por si só, uma vez que não são transmitidos em momentos fixos e regulares do dia. Os espaços comerciais não tem personalidade própria, como se verifica com os programas, para as análises incidirem neles. Para além, dos *ratings* programáticos foram identificadas algumas variáveis com capacidade de previsão desenvolvendo-se esta temática na *Questão 3*.

V.3. A ANÁLISE DOS INDICADORES ACUMULADOS DE AUDIMETRIA

Popper afirma que recorrendo apenas aos índices de audiências se torna impossível prever as tendências que correspondem ao agrado dos telespectadores, porque “tudo o que é possível recolher eventualmente, são indicações sobre as preferências dos telespectadores face aos programas que lhes são oferecidos.

Esses números não nos dizem o que devemos ou podemos propor, e esse director de cadeia também não podia saber que escolhas fariam os telespectadores perante outras propostas.” (Fernandes, 2001, p.123)

Segundo as anteriores afirmações, o conhecimento das preferências é essencial para a construção das grelhas e determinam as selecções a realizar no futuro e as actuais, existindo alguma equivalência entre ambas. Defendem que os sistemas de informação, supostamente, não permitem prever as selecções futuras e a identificação dos conteúdos a transmitir.

Na realidade os comportamentos passados, frequentemente, não apresentam uma grande capacidade explicativa e predictiva das decisões futuras (Bulck, 2006), independentemente da fonte de dados que for adoptada, e em especial quando se recorre aos inquéritos. As variáveis explicativas da exposição televisiva associadas à perspectiva de gratificações e usos do *media* também não apresentam uma significativa capacidade predictiva. A investigação mais recente demonstra que prevalece a influência estrutural, para além de uma menor influência das interacções observadas entre estas variáveis e as variáveis individuais (Cooper & Tang, 2009; Wonneberger et al., 2011; Wonneberger et al., 2012).

Os sistemas de informação são essenciais para o planeamento de *media*, uma vez que os dados devem permitir uma análise diacrónica e detalhada da exposição televisiva, com a maior objectividade possível e que só poderá ser alcançada minimizando a intervenção do telespectador. Apenas esta fonte de dados viabiliza a construção dos modelos matemáticos necessários à aquisição de publicidade televisiva. Neste capítulo pretende-se a análise dos indicadores acumulados de audimetria e demonstrar que os interesses do painel, por si só, não apresentam capacidade predictiva e explicativa da exposição ao meio televisivo. Inúmeras variáveis estruturais, também, devem ser consideradas e inviabilizam as análises de previsão considerando apenas as motivações do painel, que alguns autores defendem não ser discriminatórias dos padrões de consumo televisivo (Huertas, 2002).

A análise dos “comportamentos estruturais” (Cooper, 1996), que podem ser quantificados através dos índices de duplicação das audiências, pretende caracterizar alguns dos comportamentos que determinam a evolução dos indicadores de audimetria. Avaliamos se a audiência é variável quando se relacionam os indivíduos e os suportes ou se existem hábitos de consumo televisivo e alguma regularidade na exposição, uma vez que a estrutura dos comportamentos condiciona a viabilidade da previsão da exposição e a componente metodológica a adoptar. Pretendeu-se avaliar a predisposição para se desenvolverem determinados comportamentos de “lealdade” ou decorrentes da passividade dos telespectadores e, posteriormente, avaliar a sua repercussão ou associação com a evolução dos valores dos “ratings”, assim como as interdependências existentes.

Para além da análise da duplicação de audiências ser fundamental para o planeamento de *media*, uma vez que condiciona a selecção dos suportes publicitários, também contribui para a temática da previsão das audiências. A viabilidade de quantificação das audiências futuras encontra-se, fortemente, dependente da forma como o painel, que representa o universo televisivo, interage maioritariamente com o televisor. A análise dos comportamentos “estruturais” permite caracterizar a exposição televisiva, segundo determinadas perspectivas que avaliam o grau de regularidade na exposição e em que medida os factores estruturais a condicionam, contribuindo, até, para uma melhor selecção dos suportes a incluir em planos de *media*. Quando um anunciante compra espaço publicitário precisa conhecer a distribuição de contactos entre o seu público objectivo e os suportes a adquirir, ou seja, quantos estiveram expostos a 0, 1, 2, ... , n suportes.

Por conseguinte, é necessário determinar a duplicação das audiências entre dois suportes programáticos e comerciais e, frequentemente, recorre-se a emissões sucessivas de um mesmo programa ou a suportes transmitidos num mesmo dia. Quando se consideram as sucessivas emissões é essencial a análise dos comportamentos de lealdade programática e quando se consideram os suportes transmitidos no mesmo dia é necessária a análise dos efeitos de herança e da lealdade tipológica e ao canal.

Os resultados proporcionados pelos sistemas informáticos são, posteriormente, comparados com os que resultam da análise da exposição do painel no decorrer da campanha publicitária que poderá não ser a exposição real, mas a audimetria corresponde ao único instrumento de análise que permite com a maior objectividade possível caracterizar a exposição televisiva. Os painéis, apenas, permitem a validação dos resultados que têm que ser calculados através de modelos de previsão das audiências, que, também, devem permitir avaliar as taxas de audiências das inserções comerciais, os hábitos e o perfil das audiências dos suportes. Os estudos tradicionais que recorrem aos inquéritos não viabilizam, por motivos económicos, o cálculo das distribuições de contacto face aos suportes publicitários. As audiências dos espaços comerciais são calculadas em função das audiências programáticas, e, por vezes, recorrendo aos dois programas adjacentes, que devem apresentar duplicação de audiências e cujas audiências são calculadas por aplicação de factores correctores aos valores históricos, entre outras abordagens.

A duplicação das audiências entre as sucessivas transmissões de um mesmo programa e a sua associação com a evolução dos valores de *ratings*, neste contexto, de análise, é essencial e consta do *Apêndice F.3*. Nas análises realizadas considerou-se a totalidade da emissão diária, uma vez que na prática se pressupõe a existência de regularidade e estabilidade dos comportamentos no decorrer da tarde e da manhã, o que viabiliza a previsão dos principais indicadores de audimetria por análise das séries temporais. A exposição ao *media* não implica apenas conhecer as faixas horárias e os dias em que existe o “hábito” de ver TV, sendo a selecção apenas uma função das alternativas disponíveis em determinado momento quando a programação não é estável e das “preferências” dos telespectadores.

Quando assistem a séries, a telenovelas ou a qualquer outra tipologia que apresente uma “sequência” de conteúdos, também não existirá apenas o hábito de ver essas faixas horárias e emissões.

Na realidade as variáveis que determinam a exposição televisiva são inúmeras e nesta questão pretende-se precisamente desenvolver um modelo explicativo da exposição que coloque em evidência a complexidade de determinar os futuros comportamentos. A sequência de análises a adoptar será a que consta da seguinte *Figura V.3.1*.



Fig. V.3.1 – Análises realizadas aos comportamentos estruturais de exposição televisiva.
Fonte: Elaborado pela autora.

V.3.1. A análise dos “comportamentos estruturais” de exposição televisiva

Os modelos de duplicação das audiências correspondem a uma grande fracção dos modelos de selecção programática e são essenciais para uma maior compreensão dos padrões de consumo televisivo e para o planeamento de *media*, que depende da selecção dos suportes mais adequados à divulgação da mensagem publicitária. Com base nas análises realizadas foi viável confirmar, se a nível nacional, se verificam algumas das conclusões de trabalhos internacionais realizados neste âmbito, ainda que adoptando metodologias e um conjunto de variáveis explicativas mais abrangente e adequado à sua análise, como, por exemplo, a quantificação da duplicação através de taxas condicionadas de audiências e a sua análise recorrendo a metodologias não lineares. Em *Fernandes (2001)*, que apresenta o estado da arte nesta área, não constam quaisquer referências a análises realizadas com bases de dados nacionais e não foi viável identificar qualquer publicação nacional com uma data posterior.

Consequentemente, realiza-se uma análise comparativa face a mercados que apresentam uma diferente estruturação do mercado televisivo e de “culturas” que influenciam os padrões de consumo televisivo.

Patrick Barwise & Andrew Ehrenberg, concluíram no seu estudo de que as pessoas veem programas e não canais sendo que muitas vezes, não se tem noção em que canal se está a visionar certo programa. ... Tal como foi referido, na RTP, as pessoas seguem os seus interesses ... na SIC, consideram que a audiência é fiel aos programas e não aos canais (no entanto, diferente das restantes opiniões acredita que com o tempo acaba por se fidelizar ao canal). (*Fernandes, 2001, p.104*)

Verificámos que, sem dúvida, existe lealdade para com o canal, bem notória quando se transmitem os eventos desportivos que retiram audiência aos canais concorrentes, observando-se um fluxo de telespectadores para o canal original, imediatamente, após a sua conclusão. Excepcionalmente, observam-se efeitos de herança para os programas sucessores quando as alterações horárias não lhes permitem ver os espaços informativos que costumam seleccionar. Inúmeros eventos desportivos, transmitidos, quer nos canais privados, quer nos canais públicos, são incapazes de retirar audiência aos canais concorrentes.

Em *Fernandes (2001)*, também, é referido que apesar das pessoas verem muitos canais na semana, a lealdade ao canal é na realidade ínfima. Verificámos, com base na análise de fluxos de audiências, que os *ratings* encontram-se dependentes da duplicação de audiências num mesmo canal e dia e que, por sua vez, se encontra em maior grau condicionada pelas decisões de ligar e desligar o televisor, ou seja, pela duração do tempo disponível, do que pela exposição aos canais concorrentes. Os telespectadores que costumam realizar o “zapping” e apresentam uma forma de consumo dinâmica planeiam em menor grau, regem-se pela aleatoriedade em vez do hábito. Logo, a exposição aos canais concorrentes acaba, por vezes, por influenciar em menor grau a evolução dos *ratings*, uma vez que decorre em maior grau das transições mais breves entre os canais. Os “zappers” em geral não assistem ao programa na íntegra, quer dizer, apresentam uma reduzida fidelidade aos conteúdos. Consequentemente, existe lealdade ao canal para alguns segmentos de telespectadores e, por vezes, o nível de decisão de acesso ao *media* corresponde a ver determinado canal.

De acordo com *Fernandes (2001, p.120)*: “Dois programas emitidos no mesmo dia, consecutivos ou não, no mesmo canal, acarretam uma duplicação da audiência mais elevada (cerca de 50% ou mais), designado por efeito de herança”. Verificamos que esta afirmação, nem sempre se verifica e a duplicação de audiências em função do afastamento temporal entre episódios num mesmo dia pode ser inferior aos 20%, assim como a duplicação, nem sempre, poderá ser apenas atribuída aos efeitos de herança. A mesma autora também salienta que “... a duplicação de audiências nos programas emitidos em day time, durante a semana, pode ser ligeiramente superior ao da noite.” (*Fernandes, 2001, p.120*). Verificamos que, também, não se verifica sempre esta diferenciação, uma vez que será função das tipologias programáticas. A duplicação de audiências por vezes não é superior no decorrer do dia, uma vez que nas faixas horárias da noite pratica-se a fidelização do grande público, recorrendo à difusão do mesmo programa em faixas horários próximas e a audiência média para o total da TV é superior.

Logo, são inúmeras as incógnitas relativas à existência dos comportamentos estruturais no painel de medição das audiências televisivas nacional, respectiva interacção e influência na evolução dos *ratings*. Os modelos explicativos que apresentem uma caracterização mais abrangente dos comportamentos são essenciais para a identificação das perspectivas metodológicas mais adequadas para a construção dos modelos de exposição e inclusive para a avaliação da sua viabilidade.

V.3.1.1. A análise de fluxos de audiências e de efeitos de herança

No *Apêndice F.4* consta a análise da duplicação de audiências entre programas adjacentes ou não, mas transmitidos no mesmo dia e entre as faixas horárias. Assim, verificamos qual é a extensão horária que em geral os telespectadores apresentam o TV ligado e assistem ao mesmo canal e qual o grau de independência da audiência entre os diversos canais.

Pretendeu-se verificar como os índices de duplicação, os fluxos de audiências entre os canais concorrentes e as decisões de ligar e desligar o televisor e as variáveis de planeamento se inter-relacionam influenciando os valores dos *ratings* programáticos.

Segundo alguns autores, estes comportamentos resultam dos factores de planeamento que promovem a retenção das audiências entre programas transmitidos de forma sucessiva e apresentam independência face aos conteúdos transmitidos (*Webster, 2006*). O poder dos efeitos de herança entre programas fundamenta-se no pressuposto de que os telespectadores assistem ao *media* de forma “passiva”, vendo determinado programa, simplesmente, porque sucede uma anterior selecção ou antecede um programa que pretendem ver.

Os efeitos de herança correspondem a um dos comportamentos de duplicação de audiências entre suportes publicitários adjacentes e que transcende a lealdade ao canal. Têm sido, essencialmente, atribuídos a factores de inércia; aos factores de planeamento que tornam o início da programação dos diversos canais desfasados no decorrer do tempo e à tendência dos telespectadores para esperar pelo início de determinado programa. Pretendeu-se responder, essencialmente, às seguintes questões que visam caracterizar e explicar a duplicação das audiências e correspondente repercussão na evolução dos *ratings*:

➤ Qual a influência da contra programação na evolução da duplicação primária, ou seja, no indicador que quantifica a audiência que transita entre programas adjacentes ou próximos, recorrendo a taxas condicionadas de audiências?¹ Quais são os canais que penalizam ou favorecem cada um dos canais generalistas em maior grau, melhor dizendo, qual é o grau independência entre a audiência e os *ratings* de cada canal?

¹ Este indicador corresponde a uma tipologia de dados minoritariamente utilizada na bibliografia internacional que em maior grau opta por comparar os valores de ratings de cada par de programas ou as taxas de audiências conjuntas.¹

- A correlação entre os valores de *ratings* de programas adjacentes traduz o grau da duplicação primária existente? Os *ratings* são influenciados em maior grau pelos padrões de disponibilidade e de predisposição para ver a TV ou por parte da duplicação das audiências e dos fluxos de audiências entre os canais?
- Os factores de planeamento condicionam a evolução da duplicação primária das audiências? O canal, as faixas horárias, as tipologias programáticas, as sinergias tipológicas, ou seja, a coincidência entre as tipologias de cada par de programas, o desempenho da programação adjacente influenciam a ordem de grandeza e a variabilidade observada na duplicação primária das audiências?

O reduzido número de publicações que foi viável identificar, nesta área, incide em dados Norte Americanos, e apenas uma publicação recorre a dados posteriores ao ano 2000. Segundo *Webster (2006)*, nas análises realizadas, em 1985, constatou-se que o factor que influenciava em maior grau a duplicação das audiências correspondia ao número de possibilidades de transição entre programas adjacentes, mas, recorrendo a bases de dados mais recentes, destaca-se as sinergias tipológicas. Justifica esta alteração, com o acréscimo observado no número de canais.

Este autor salienta que, quando os telespectadores apresentavam poucas opções ao nível dos canais, o número de alternativas disponíveis era um factor relevante, mas, actualmente, perante a diversidade de canais disponíveis, facilmente, existem opções alternativas. Consequentemente, num contexto de uma maior diversidade de opções este factor tende a apresentar uma reduzida capacidade explicativa. Ambas as variáveis estão negativamente relacionadas com a duplicação de audiências, apresentando um maior coeficiente a exposição aos quatro principais canais, em detrimento da exposição à totalidade dos canais existentes. Considerando um modelo de Regressão linear múltipla, a variável explicativa mais relevante, segundo o autor, corresponde às sinergias tipológicas, seguindo-se o *rating* do programa sucessor, com um coeficiente beta positivo, o número de opções disponíveis nos quatro principais canais com um coeficiente negativo e o *rating* de “Lead_In”, com um valor negativo.

Este modelo de Regressão linear apresenta um coeficiente de determinação no valor de 0,58. Incluir no modelo, os canais com menores níveis de exposição e os canais por TV Cabo nada acrescenta, apenas uma aparente aleatoriedade. O autor refere que as tipologias apresentam uma relevância que deve ser considerada como sendo relativa, uma vez que representam um vasto conjunto de programas perdendo-se no seu conjunto particularidades dos mesmos que podem ser relevantes.

Identificámos perspectivas contraditórias a respeito da repercussão dos conteúdos e das tipologias programáticas, por exemplo, segundo *Webster e Wakshlag (1983)*, *Webster (1985)*, *Davis e Walker (1990)* verifica-se uma duplicação de audiências prevista por semelhança nas tipologias transmitidas e para *Goodhardt, Ehrenberg e Collins (1975)* apenas os factores de planeamento condicionam os padrões observados. *Headen, Klompmaker e Rust (1979)* salientam que os comportamentos observados nos USA apresentam uma maior complexidade face aos trabalhos Britânicos e que as tipologias apresentam uma ligeira repercussão na previsão da duplicação das audiências.

Com base nas análises realizadas a estes factores, entre muitos outros, que foram avaliados recorrendo, não só, aos programas adjacentes dos períodos de maiores audiências, mas, sim, considerando a totalidade da emissão diária, e por aplicação de metodologias de Regressão linear e não linear verificamos que os resultados a nível nacional divergem.

Considerando as mesmas faixas horárias e a tipologia de dados que se encontram descritos em *Webster (2006)*, do ponto de vista linear, constatamos que a variável mais relevante não corresponde às sinergias de tipologias entre os programas adjacentes, mas, sim, ao número de telespectadores que mantêm o televisor ligado na interface de dois programas, sendo esta variável responsável por 60% de variância explicada. Incluindo a variável TV_OFF, que representa a percentagem de telespectadores que desliga a TV, a percentagem de variância explicada é de 70%, superando o valor do coeficiente de determinação do modelo de Regressão linear apresentado em *Webster (2006)*.

Consequentemente, os factores de disponibilidade e de predisposição para assistir ao *media* condicionam, maioritariamente, a duplicação das audiências, e podem ser, frequentemente, independentes dos conteúdos transmitidos, imperando a “experiência televisiva”. Os programas adjacentes ou próximos na grelha de programas podem partilhar uma maior percentagem de telespectadores, porque um número mais elevado de pessoas tende a estar disponível e a ver TV em períodos de tempo próximos.

O modelo apresentado por Webster e Wakshlag (1983) traduz a existência de padrões de consumo televisivo correlacionados que podem ser atribuídos aos factores de disponibilidade, frequentemente, independentes dos conteúdos acessíveis em determinado momento. Os padrões de exposição ao *media* são responsáveis pela existência de fluxos de audiências estruturalmente bem definidos. A duplicação das audiências deverá ser atribuída no decorrer de uma sessão televisiva à disponibilidade da audiência, que por sua vez tende a assistir a um determinado conjunto de sucessivos programas. Existindo um padrão de utilização do *media*, que poderá ser analisado recorrendo às taxas médias de audiências para o total da TV, podemos atribuir a este factor os efeitos de herança observados.

Segundo as análises realizadas à variável de audimetria *Tvr%*, verificamos que a variabilidade observada neste indicador no decorrer dos vários anos e para um idêntico horizonte temporal, não transcende significativamente os valores que resultam da análise da sua variabilidade no decorrer de alguns meses onde se mantém a estruturação da grelha de programas, reforçando o pressuposto da existência de alguma independência entre os conteúdos e a exposição ao *media* para parte do painel. As taxas médias de audiências apresentam, no geral, valores que são moderados ou reduzidos, ou seja, uma menor percentagem do painel costuma assistir ao *media* e numa percentagem ainda menor assistem com uma maior regularidade, atendendo aos resultados das análises de duplicação das audiências.

Em relação à repercussão da programação concorrente nos índices de duplicação, para os dados do canal SIC verificamos que a exposição ao canal RTP2 está positivamente

correlacionada, enquanto a exposição ao canal RTP1 e ao canal TVI apresenta coeficientes de correlação negativos, mas que, frequentemente, pouco acrescentam ao nível do coeficiente de determinação. A exposição ao vídeo, aos canais por TV Cabo e aos canais por satélite não apresenta coeficientes estatisticamente significativos.

Podemos concluir que, analogamente, ao que se verifica nos USA, a exposição aos canais concorrentes não apresenta uma grande repercussão na evolução da duplicação primária das audiências. Segundo *Webster (2006)*, existe uma selectividade “defacto”, ou seja, os telespectadores perante uma grande variedade de opções restringem as suas selecções a um reduzido número de canais para facilitar o processo de selecção. Por outro lado, o painel nacional, na sua maioria, apenas tem acesso aos quatro canais generalistas, por conseguinte a exposição aos canais por cabo não deveria influenciar de forma significativa a evolução da duplicação primárias das audiências. A exposição aos canais generalistas, apesar de existir, não se repercute de forma muito significativa na duplicação das audiências dos principais canais.

Verificamos que uma significativa percentagem da audiência que assiste ao espaço informativo de determinado canal opta por desligar a TV após ter assistido a alguns programas posteriores e numa menor percentagem transita para outro canal. Os telespectadores que ligam o televisor no início do horário nobre correspondem aos que direccionam uma maior “fidelidade” a determinado canal. Quando a sessão se inicia em períodos posteriores e os telespectadores pretendem a selecção de determinados conteúdos ou tipologias de programas, as selecções podem incidir em mais do que um canal quando a sessão abrange mais de um programa, existindo um maior fluxo de audiências entre os dois canais privados generalistas. Segundo *Meyer e Muthaly (2008)*, os telespectadores das faixas horárias mais tardias apresentam uma menor lealdade ao canal de curto e médio prazo e os segmentos sociodemográficos que maximizam estes indicadores correspondem às faixas etárias superiores aos 50 anos, lares com apenas um telespectador, classe sociais intermédias e os telespectadores que assistem maioritariamente aos principais canais generalistas. Quem apresenta um maior consumo televisivo, também, apresenta uma maior lealdade, mas apenas de curto prazo.

Logo, os valores mais elevados de *ratings* estão associados a uma maior lealdade de curto prazo e que foi confirmada recorrendo às análises dos fluxos de audiências entre canais. Quem apresenta uma maior exposição televisiva, também, transita em menor grau entre os diversos canais, existindo um maior envolvimento programático.

Outros factores, também, se repercutem na duplicação das audiências, entre os quais o desempenho da programação adjacente. Nas análises realizadas à evolução da duplicação de audiências verificamos que o *rating* de “Lead_In”, ou seja, do programa que antecede o programa em análise apresenta um coeficiente de correlação irrisório e o valor do *rating* do programa sucessor apresenta um valor moderado e positivo. Atendendo a estes resultados, um programa com um elevado valor de *rating* não implica necessariamente que se observem elevados valores de duplicação primária na sua sequência (“... the large audience of a highly rated lead-in is typically lost in disproportionate numbers, resulting in low levels of primary duplication” (Webster, 2006, p.335)), mas maiores valores de *ratings* sucessores estão associados a um maior fluxo de audiência vindo do programa adjacente e anterior.

Relativamente à associação da variável em estudo com os valores dos *ratings*, o maior valor corresponde ao *rating* parcial do programa de “Lead_Out” ($\beta=0,55$, $R^2=0,30$), assim como se verifica em Webster (2006). Adicionando o *rating* parcial do programa de referência ou de “Lead_In” ($\beta=-0,52$, $R^2=0,43$) confirma-se o valor negativo apresentado pelo autor. A associação linear com o valor sucessor é positiva e com o valor antecessor é negativa. Não existe associação se considerarmos os valores globais de *ratings* dos programas adjacentes, ou seja, considerando a duração global do programa em vez de apenas uma parte programática. Uma vez que impera a influência dos factores de disponibilidade e de predisposição para assistir ao *media*, quantificados através das variáveis de TV_ON e TV_OFF, após um programa com um elevado *rating* não se verifica necessariamente que grande parte da sua audiência seja transportada por inércia para o programa subsequente. Verificámos, recorrendo aos índices de lealdade programática, que apenas no máximo para metade do painel que assiste a determinado programa existe uma exposição mais regular ao *media* no decorrer do tempo.

Para os telespectadores que apresentam uma reduzida lealdade para com o meio televisivo, as selecções podem ser mais pontuais, e apenas para assistir aos seus programas “favoritos” abrangendo um curto espaço de tempo e desligando o televisor após a sua conclusão. Um programa sucessor que apresente um maior valor de *rating* poderá implicar a existência de algum fluxo de audiências vindo do seu programa predecessor.

Em relação à questão levantada em Webster (2006) quanto a se o grau de associação linear entre os *ratings* permite avaliar a ordem de grandeza da duplicação das audiências, verificamos que os valores obtidos nas análises realizadas diferem. Segundo o autor, não existem diferenças nos coeficientes de correlação de *ratings* quando se consideram programas com elevada ou baixa duplicação, apresentando o valor de 0.80. Segundo as análises realizadas, considerando valores parciais de *ratings* e reduzidos valores de duplicação, o coeficiente de correlação linear é de 0,87 e em contextos de elevada duplicação de -0,18. Confirma-se uma elevada correlação entre *ratings* nos contextos em que a duplicação primária não ultrapassa os 60%. Sendo o valor de correlação negativo para valores de duplicação superiores aos 60%, podemos concluir que os maiores valores de *rating* de um programa são acompanhados por menores valores no programa adjacente. Esta tendência poderá corresponder a faixas horárias em que se observa uma tendência ascendente ou descendente na evolução dos valores dos *ratings*, consequentemente existe uma associação inversamente proporcional entre os indicadores de audimetria.

Webster (2006) levanta a seguinte questão: Se existir uma elevada correlação entre os valores de *ratings* de programas que partilham poucos telespectadores o que poderá explicar esta tendência? Uma elevada correlação significa que os valores variam proporcionalmente e de forma linear, existe uma recta que estabelece a relação funcional entre as variáveis. Um acréscimo de uma unidade em um dos indicadores implica igual acréscimo na variável dependente. Os telespectadores variam ao nível do tempo de exposição entre programas adjacentes, influenciando os valores de *ratings*.

Ainda que a duplicação de audiências fosse de 100%, os *ratings* não seriam necessariamente iguais, tudo dependerá do tempo de exposição e das decisões de desligar a TV ou mudar de canal.

Os *ratings*, também, são influenciados pela exposição dos telespectadores que decorre a partir de quem assiste maioritariamente aos canais concorrentes. Com o controle remoto, facilmente, um telespectador poderá ligar determinado canal para ver o que está a ser transmitido nesse momento, ainda que a maior parte do tempo da sua sessão televisiva esteja associada a canais alternativos que façam parte do seu “repertório”.

Com base nos dados analisados e que se referem a períodos após as 19H, os programas que herdaram no máximo 60% dos telespectadores do programa anterior, apresentam uma elevada associação linear de valores de *ratings*. Os programas que herdaram mais de 60% da audiência, por sua vez, apresentam valores muito mais reduzidos do ponto de vista linear e apresentam um coeficiente negativo. Todavia, a associação poderá ser significativa, mas, não, linear. Webster (2006) refere que as audiências se têm tornado cada vez mais fragmentadas e a redução ao nível da dimensão dos *ratings*, também, reduz a variabilidade dos valores de shares.

Verificámos, com base nas análises realizadas, que os *ratings* em geral não são muito elevados, mas a sua variabilidade, sim. Quando os seus valores são calculados para programas, a variabilidade pode ser muito significativa, e ainda superior se considerarmos os valores de *shares* de audiências. Verificámos, também, que, no decorrer da noite, os *ratings* vão decrescendo em valor, assim como a sua variabilidade. Webster (2006) refere que os *ratings* são reduzidos, como consequência dos programas se dirigirem a um cada vez menor número de nichos de mercado e à selecção “defacto”. Esta realidade levaria a uma maior correlação entre as observações adjacentes. Em Portugal, verificamos que existe alguma independência entre os públicos de cada canal para uma grelha de programas já estabelecida, resultando a ordem de grandeza dos efeitos de herança em maior grau das decisões de ligar e desligar o TV.

Quando uma grelha de programação é reestruturada, em função dos segmentos, poderá existir uma alteração, ou, não, no conjunto de canais e faixas horárias que são seleccionados com uma maior frequência. Constatámos que existe uma maior associação linear entre os *ratings* de um programa e do seu programa sucessor.

A eventual atractividade exercida por parte de determinados programas poderá não incentivar de forma muito pronunciada a uma prévia ligação do televisor, o que pode ser comprovado com a análise da duplicação primária das audiências. Após assistirem ao programa que pretendem, independentemente dos factores subjacentes, os telespectadores são transportados por inércia para os programas subsequentes, mas grande parte do painel opta por desligar o televisor, e, por vezes, talvez, por mudar de canal, sendo a sua ordem de grandeza função das faixas horárias consideradas.

Atendendo aos valores dos coeficientes de correlação entre programas adjacentes, a aplicação em análises de previsão das Regressões Locais, sejam estas lineares ou não, deveria ser uma boa opção, apesar de ser superior a associação linear face aos períodos subsequentes. Contudo, as análises de previsão de $Tvr\%$, recorrendo a estas metodologias, também, não conduziram a resultados satisfatórios. Considerando os valores de *ratings* referentes aos canais, ou seja, a períodos horários, a TVI maximiza os valores de correlação entre as faixas horárias adjacentes, seguindo-se, a SIC e a RTP1, sendo estes valores superiores face aos valores programáticos. O facto destes valores serem superiores aos programáticos indicia que face a períodos horários de menor duração (considerou-se 30 minutos nas análises realizadas) a duplicação de audiências é superior. Consequentemente, não parece existir uma forte fidelização das audiências, o que será confirmado com a posterior análise dos indicadores sociodemográficos de audimetria.

Também, foi avaliada a influência, na duplicação das audiências, de outros factores que não constam das análises realizadas em Webster (2006). Por exemplo, verificamos que a duplicação primária não é superior quando cada par considerado de partes programáticas corresponde ao mesmo programa.

Considerando a tipologia do programa de referência de Grau_1, verificamos que os programas sucessores do espaço informativo, maximizam a duplicação de audiências. Para o período em análise, o grau de fidelidade dirigido aos espaços informativos não condiciona a ordem de grandeza dos efeitos de herança nos períodos adjacentes. Uma maior fidelidade não implica necessariamente a existência de maiores níveis de audiência em programas adjacentes.

Verificamos que entre as tipologias que se associam em maior grau aos programas que constam das grelhas nacionais, como, por exemplo, a informação, o divertimento e a ficção não se observam diferenças nas médias associadas à duplicação primária das audiências e apresentam um valor médio de 60%. O desporto para além de minimizar estes valores também apresenta uma grande variabilidade. Além do mais, as tipologias programáticas não devem ser consideradas como um factor chave de previsão dos comportamentos de consumo televisivo, uma vez que o desempenho de um programa não implica que outro programa de semelhantes características, também, venha a alcançar a mesma ou uma exposição que lhe seja equivalente.

A duplicação de audiências, durante a semana, é superior face ao que se verifica ao fim de semana (66% vs. 59% de valor médio), uma vez que o tempo de contexto difere, sendo a exposição televisiva mais irregular em dias de lazer se considerarmos as faixas horárias que não correspondem ao final do dia. Considerando o período horário, apenas a noite se diferencia com os menores valores observados (52% de valor médio), ou seja, entre as 22H e as 24H, e no pré prime time e no horário nobre as médias observadas coincidem nos 70%.

Em suma e de acordo com *Webster (2006)*, não existem indícios de estarem a decrescer os efeitos de herança e, em Portugal, verificamos que em função das faixas horárias estes índices ainda podem ser consideravelmente razoáveis, sendo relevante uma estruturação da grelha que permita tirar o maior partido destes comportamentos.

V.3.1.2. A análise da lealdade programática

Uma segunda vertente de análise da duplicação das audiências corresponde à lealdade programática. No *Apêndice F.3*, constam alguns resultados que traduzem os níveis de duplicação de audiências entre diferentes episódios ou transmissões de um mesmo programa ao longo do tempo. As análises de lealdade programática, que constam da bibliografia internacional da especialidade, têm incidido, essencialmente, na análise da duplicação de audiências entre os suportes programáticos, sendo consequentemente, também, denominadas de lealdade programática, apresentando no Reino Unido, valores que não transcendem os 55% e nos USA valores inferiores aos 25% apesar de uma maior regularidade observada por vezes na evolução dos valores de *ratings*. Os reduzidos valores observados de duplicação das audiências têm sido atribuídos aos factores de planeamento, a factores de disponibilidade e a um reduzido envolvimento com o *media*.

Contrastando com a literatura da especialidade recorreu-se às taxas de audiências condicionadas e procurou-se responder a algumas questões que visam a caracterização destes comportamentos a nível nacional e a avaliação das respectivas implicações para o planeamento de *media*.

Verificámos se os comportamentos de “repeat viewing” estão presentes no painel nacional de forma equiparável ao que se verifica a nível internacional. Constatou-se em geral a existência de reduzidos a moderados valores destes índices para diferentes episódios ou transmissões de um programa ou espaço programático, mas podendo, por vezes, alcançar e ultrapassar os 50% para determinadas tipologias programáticas ou faixas horárias. Apresentam certamente valores superiores aos que constam da bibliografia internacional. Em Portugal, para determinadas faixas horárias e tipologias existe uma maior regularidade na exposição face ao que se verifica em países mais desenvolvidos como os USA (*Ehrenberg & Wakshlag, 1987*), Reino Unido (*Goodhardt, Ehrenberg & Collins, 1987*) ou na Alemanha (*Zubayr, 1999*).

A nível internacional, nos anos 50-60, observavam-se valores de duplicação de audiências no intervalo 50%-60%, mas, a nível nacional, para o período em análise podem, por vezes, ser observados valores que se aproximam dos 50%. Na Alemanha, no final da década de 90, aproximavam-se dos 25%-35% em média. Em Portugal, pelo menos entre 40% a metade do painel que assiste a determinado programa apresenta, nos períodos de maiores audiências, uma exposição mais regular.

A quantificação das audiências para uma sequência de suportes poderá ser, eventualmente, realizada em função da dimensão das suas audiências conjuntas. Por conseguinte, foi necessário avaliar a seguinte questão: *“Qual a repercussão da lealdade programática nos valores de ratings que são essenciais para o planeamento de media? Qual a influência do distanciamento temporal entre as transmissões nos índices de duplicação?”*

Uma vez que os índices de duplicação ficam em geral aquém dos 50%, se os *ratings* resultassem apenas da audiência herdada de episódios anteriores os valores dos *ratings* seriam francamente mais variáveis do que são na realidade para determinados períodos, ainda que os *ratings* correspondam a taxas médias de audiências ponderando os telespectadores em função do tempo de exposição. Um programa pode “herdar” em média uma determinada percentagem da audiência de um episódio anterior e os *ratings* ponderam-nos em função do tempo de exposição ao evento, consequentemente, ainda que a duplicação fosse de 100% e não existissem novos telespectadores a assistir ao programa, os valores de *ratings* não seriam necessariamente equivalentes.

A audiência de um programa inclui não só parte da audiência do episódio anterior, como também, de novos telespectadores que podem ter assistido a episódios mais afastados no tempo. Porém, segundo as análises realizadas, um maior desfasamento temporal entre os episódios nem sempre apresenta uma forte repercussão na duplicação das audiências, observando-se um decréscimo de apenas um dígito percentual nos valores destes índices.

Foi necessário avaliar qual é a influência do canal e das alterações horárias na duplicação de audiências e se os índices de duplicação resultam de uma exposição activa ou passiva. A estruturação da grelha de programação recorrendo a programas em séries visa do ponto de vista da indústria televisiva facilitar a previsão de alguns indicadores, levando-os por vezes a nível internacional a apresentar uma considerável estabilidade de *ratings*, segundo alguns autores. Para os anunciantes, esta tipologia de programas deveria tornar o consumo televisivo e, conseqüentemente, o resultado das campanhas comerciais mais facilmente previsíveis permitindo-os otimizar o número de exposições aos segmentos-alvo pretendidos.

As frequentes alterações horárias programáticas, em alguns dos anos em análise, contribuem para anular este efeito, uma vez que os programas que são adiados para faixas horárias mais tardias não conseguem reter grande parte da audiência. A popularidade de um programa influencia esta capacidade, sendo notória ao fim de semana onde a variabilidade nos valores dos *ratings* apresenta menores valores face aos valores semanais.

Em geral na literatura da especialidade, a análise destes índices incide nos programas que são transmitidos regularmente e com uma considerável estabilidade ao nível dos horários de transmissão. No entanto, para determinados períodos em análise as telenovelas e as séries da noite, em especial os programas transmitidos, na TVI, variam significativamente nos seus horários, contribuindo para uma considerável variabilidade dos valores de *Rat_Prog%* ou de *Rat_Período%*, ou seja, para o canal e para uma redução nos valores dos índices de lealdade programática.

As análises realizadas permitiram avaliar se no painel predominam os comportamentos de consumo televisivo mais activos ou ritualizados e com um cariz de maior passividade. Em função das faixas horárias, tipologias e dias os resultados variam, uma vez que estas variáveis influenciam a composição sociodemográfica do painel que condiciona a estrutura dos comportamentos de consumo televisivo e as práticas quotidianas dos indivíduos.

Os índices de duplicação não transcendem para o período em análise os 50%, ou seja, sensivelmente metade do painel que assiste a determinado programa. Os valores dos *ratings* que representam a exposição ao público alvo também são moderados a reduzidos.

Consequentemente, grande parte do painel não apresenta regularidade na exposição e quando o faz em função dos contextos e das características dos segmentos poderá ser por uma procura activa de conteúdos ou prevalecer a experiência televisiva. Os telespectadores que assistem, frequentemente, aos mesmos programas são considerados como pouco activos nas suas selecções. Quem assiste à TV independentemente da programação são designados na literatura da especialidade como orientados para o *media* e os restantes como orientados para os conteúdos.

Aos primeiros, associa-se um comportamento passivo, ritualizado e segundo hábitos adquiridos, enquanto os segundos ao procurar ver determinados programas apresentam um comportamento mais activo e instrumental. No segmento que apresenta determinados hábitos de consumo televisivo podem estar presentes ambos os comportamentos.

Os moderados índices de duplicação observados no decorrer de grande parte da emissão diária podem resultar de alterações nos padrões de disponibilidade e de predisposição do público, como também das oscilações horárias que se verificam, frequentemente, nos programas, levando a elevados valores de variabilidade de *Tvr_Prog%* e que especialmente em dias laborais não incentivam à permanência das audiências. A forte repercussão horária leva-nos a concluir que impera o segmento orientado para o *media*. Concordamos que a audiência apresenta maioritariamente um comportamento activo em ambientes estruturais ou mais passivo. Inúmeros factores estruturais influenciam a duplicação das audiências, nomeadamente as faixas horárias, as tipologias programáticas, a popularidade do programa e forma de transmissão. Foram realizadas análises no sentido de avaliar a repercussão das variáveis mencionadas.

Segundo as análises realizadas às taxas médias de audiências para o total da TV, e aos *ratings* diferentes tipos de comportamentos parecem poder ser associados aos canais estatais versus privados, mas do ponto de vista dos índices de duplicação a nível médio não se observam diferenças em função dos canais considerados. Contudo, se consideramos determinadas tipologias programáticas existe alguma diferenciação. Em princípio, podemos associar motivações instrumentais e activas em maior grau ao canal estatal, para o período em análise. O comportamento passivo será o que depende maioritariamente da disponibilidade dos telespectadores e do seu tempo vazio e em menor grau dos conteúdos, enquanto o comportamento activo será o motivado por ver determinados programas.

Apesar do canal estatal apresentar para grande parte do painel a procura de “informação”, os factores de disponibilidade, enquanto factor estrutural, também, imperam podendo as alterações horárias incutir uma grande variabilidade nos indicadores de audimetria. No entanto, também, verificámos, que excluindo os programas de maior notoriedade este canal apresenta uma maior estabilidade de *ratings* interanual e bimensal e de taxas médias de audiências para o total da TV face aos canais privados generalistas. Também, ficou demonstrado que a evolução agregada não traduz o que se verifica a níveis mais desagregados de análise. A existência de motivações instrumentais, ainda que condicionada por factores de disponibilidade a nível nacional, incutem uma maior regularidade nos indicadores de audimetria face a uma grelha de programação que não incluía eventos desportivos e outros programas de maior notoriedade. Os comportamentos de lealdade programática e os efeitos de herança, em geral, não são explicados a partir do conteúdo programático envolvido, estando associados a comportamentos de maior passividade dos telespectadores. A duplicação das audiências associada à lealdade programática na realidade pode resultar de dois contextos:

(1) Os telespectadores podem assistir ao mesmo programa por apresentarem o televisor ligado com alguma regularidade nessas faixas horárias. Uma vez escolhido o canal, o consumo passa a ser ritualizado.

Verificamos que, durante a semana, esta parece ser a tendência observada em maior grau e em especial para os canais privados. No entanto, os espaços informativos da noite da RTP1 que apresentam elevados *ratings* e taxas médias para o total da TV não apresentam índices de lealdade programática que superem significativamente os 40%, apesar da possível procura dos conteúdos informativos. Os factores de disponibilidade imperam, e a este canal foi possível associar as maiores faixas etárias e classes sociais. Do ponto de vista das faixas etárias tem-se associado motivações ritualísticas em maior grau aos segmentos com mais idade, mas os índices de duplicação não são superiores aos dos canais generalistas privados.

(2) O motivo para ligar a TV estará relacionado com o conteúdo e não com a faixa horária. Verificamos, por exemplo, que a principal telenovela da noite da SIC, em alguns dos anos, por vezes, apresenta valores de *ratings* superiores ao do espaço informativo da noite, levando a pressupor que existem telespectadores que ligam o televisor apenas para ver este programa. Verificamos que as alterações horárias observadas nos espaços informativos da tarde e da noite, quer sejam semanais ou de fim-de-semana, levam a alterações por vezes muito significativas nos valores do *ratings*. Apesar da possível procura dos conteúdos informativos, os factores de disponibilidade imperam e condicionam a exposição televisiva. Com base nas análises realizadas os resultados parecem confirmar as conclusões internacionais de uma maior passividade do público, ou pelo menos de uma forte influência das variáveis estruturais, o que não implica inevitavelmente uma maioritária regularidade na exposição.

Foram realizadas algumas análises para verificar se os índices de lealdade programática encontram-se influenciados em maior grau pelos factores de disponibilidade ou pelos fluxos de audiências entre canais. Ao nível dos efeitos de herança verificámos que as decisões de ligar e desligar o televisor são as variáveis que condicionam a evolução destas variáveis, e em menor grau os fluxos de audiências para os canais concorrentes. Logo, ao nível dos índices de lealdade programática, o público que não selecciona programas consecutivos deve-se certamente em maior grau a factores de disponibilidade, e não à selecção de outros programas de canais concorrentes.

A nível internacional, o decréscimo na duplicação de audiências atribui-se em maior grau aos factores de disponibilidade do painel e não à atractividade exercida por parte da programação concorrente. Segundo alguns autores, o mundo do trabalho tem-se transformado significativamente, desde as iniciais rotinas estandardizadas, até hoje em dia, se verificar a existência de horários de trabalhos mais flexíveis e trabalho em part-time etc., especialmente em países mais industrializados ou pós-industrializados, como os USA e a Alemanha. A dissolução das anteriores rotinas de trabalhos acaba por se repercutir nas rotinas das actividades lúdicas, ou seja, na disponibilidade para se poder assistir regularmente a determinado programa.

Uma fracção do painel nacional, por vezes, apresenta rotinas de consumo nas faixas horárias de maiores audiências, o que é significativo atendendo à tendência existente a nível internacional. Os menores índices de duplicação observados em alguns contextos podem resultar de alterações nos padrões de disponibilidade do público, agravados pelas oscilações horárias que se verificam, frequentemente, nos programas levando a elevados valores de variabilidade de *Tvr_Prog%* e que especialmente em dias laborais não incentivam à permanência das audiências.

Persistem inúmeras dúvidas a respeito da influência das tipologias programáticas, dos *ratings* e das faixas horárias nos índices de lealdade programática. De acordo com *Fernandes (2001)*, nas referências bibliográficas identificadas, os autores, salientam que a tipologia do programa não contribui para a duplicação das audiências. Segundo *Meyer e Muthaly (2008)*, as tipologias que apresentam uma maior lealdade no seu decorrer e a longo prazo correspondem aos espaços informativos e às magazines, enquanto os filmes e as comédias minimizam estes indicadores. As correlações costumam ser mais elevadas quando se recorre à lealdade de longo prazo. Ao nível da influência das tipologias e horários de transmissão verificamos que estas questões na literatura da especialidade não foram claramente respondidas. Às telenovelas, em geral, associam-se maiores valores e nas análises realizadas, também, o verificamos, sendo os valores superiores no decorrer da noite face ao período de pré prime time.

A influência dos *ratings* nestes indicadores, segundo alguns autores, apresenta alguma lógica uma vez que os índices de duplicação têm que ser superiores quando os *ratings*, também, o são. Perante um programa que apresente uma maior audiência é obvio que, também, será maior a probabilidade de se manterem em episódios subsequentes maior número de telespectadores.

Relativamente à duplicação das audiências entre transmissões de um mesmo programa, *Fernandes (2001, p.120)* refere que: “... existe também o denominado “double jeopardy effect”, que se verifica quando uma estação de televisão tem uma fraca audiência e espectadores pouco leais ou regulares.” Verificamos que nas faixas horárias mais tardias, apesar da menor dimensão dos valores dos *ratings*, a duplicação de audiências pode ser considerável para alguns programas. Por vezes, é referido que não existem audiências diminutas, mas, sim, mais leais aos programas. Nas análises realizadas verificamos que esta constatação nem sempre é válida. Em relação à obrigatoriedade de maiores valores de *ratings* levarem a maiores valores de duplicação, também, nem sempre se verifica. Os *ratings* correspondem a taxas médias de audiências, e, logo, o mesmo valor pode resultar de um menor número de telespectadores que assistem a uma maior duração do programa ou de um número mais elevado de telespectadores que assistem pouco tempo ao programa. No primeiro contexto a duplicação de audiências deverá ser superior.

Relativamente à influência do número de canais nos índices de lealdade programática, verificamos que um acréscimo no número de canais pode levar a um decréscimo nestes índices. No entanto, a nível nacional, o número de lares com TV Cabo ainda não é muito elevado para o período em análise e por isso verifica-se a sua menor representatividade no painel de audimetria. Os segmentos que minimizam as taxas médias de audiências para o total da TV são a classe social AB e os Lares com TV Cabo. Sendo estes os segmentos que apresentam menores taxas médias de audiências para o total da TV, também, não podem contribuir significativamente para os menores valores de lealdade programática.

Coloca-se a questão de se os índices de lealdade programática podem contribuir para a previsão dos *ratings*. Do ponto de vista da previsão dos *ratings*, em função das faixas horárias uma vez que a duplicação de audiências será no máximo de 50% considerando metade da duração dos programas e os valores dos *ratings* são reduzidos, apenas para uma menor percentagem do painel em determinados contextos se poderá prever as suas selecções com base no seu historial de consumo televisivo para idênticas faixas horárias. A maior parte do painel não estará a assistir à TV em faixas horárias equivalentes.

A previsão dos *ratings* está fortemente condicionada pelos factores de disponibilidade que são demasiado difíceis de prever. Não será viável prever todas as condicionantes laborais ou de outro género a cada telespectador está sujeito e sobre as quais grande parte das vezes não tem poder de decisão.

Grande parte das análises de previsão que existem na literatura da especialidade, apenas consideram os factores temporais, o canal e as tipologias programáticas ignorando os restantes factores que contribuem para a evolução dos principais indicadores de audimetria. Na actualidade, a nível nacional, estas abordagens ainda podem produzir resultados satisfatórios, como foi constatado, atendendo ao que se verifica a nível internacional. No entanto, em princípio, a médio prazo a tendência será para cada vez mais este tipo de abordagens produzir resultados insatisfatórios. Porém, nos contextos em que sejam viáveis as análises de Médias Móveis ponderadas, estas podem viabilizar a previsão recorrendo à duplicação das audiências.

Relativamente à questão de quais são os factores que podem levar ao desenvolvimento deste tipo de comportamento, por vezes, é salientado a previsibilidade dos conteúdos transmitidos e uma maior familiarização com as personagens. No entanto, existem factores estruturais que nitidamente influenciam este indicador, nomeadamente as faixas horárias, as tipologias programáticas, a popularidade do programa e a forma de transmissão.

Utilizando maiores desfasamentos temporais entre episódios, por forma, a atenuar a repercussão da possível existência de rotinas de trabalho verificamos que os índices, em geral, mantêm a mesma ordem de grandeza face aos programas consecutivos e, por vezes, inferiores quanto mais afastados estiverem os programas em análise.

No que diz respeito à influência das variáveis sociodemográficas na duplicação de audiências, será de esperar que quem apresenta uma maior exposição ao meio televisivo também deverá maximizar este indicador. Os segmentos com maiores valores das taxas médias de audiências para o total da TV e de *ratings* foram identificados, nos restantes capítulos, para cada um dos canais em análise.

Uma vez que, em geral, as alterações horárias dos programas incutem uma significativa variabilidade nos valores de *ratings* e de taxas médias de audiências para o total da TV, aparentemente para grande parte do painel existirá uma exposição temporal cíclica e até certo ponto independente do conteúdo transmitido. No *Apêndice F.5* constam determinadas análises que são necessárias para a construção dos planos de *media*, como, por exemplo, como proceder à selecção dos suportes que será necessário em determinadas etapas deste processo uma vez que são inúmeras as combinações possíveis de suportes.

No *capítulo VI* descreve-se uma proposta de construção dos planos recorrendo aos dados dos ficheiros de base de audimetria, descrevendo-se as diversas etapas que a constituem e que transcende a necessidade de previsão dos principais índices de audiências. Para além das análises de previsão da exposição agregada ou de determinados indicadores, a construção dos planos de *media*, também, envolve a necessidade de segmentar o painel e as faixas horárias e de “optimizar” a selecção dos suportes e dos públicos objectivo. Os suportes a incluir nos planos devem apresentar a maior e “melhor” exposição aos segmentos-alvo da campanha publicitária em função de determinadas restrições.

Ao invés de adoptar o conceito de frequência efectiva, os suportes devem ser seleccionados de modo a obter a “melhor exposição” e que será a que for mais adequada a determinado segmento de telespectadores e às características da campanha.

V.3.1.3. A análise da lealdade tipológica e ao canal

Nas análises que constam do *Apêndice F.5*, promove-se a associação de programas, tipologias e canais aos indicadores sociodemográficos que constam das bases de dados de audimetria. Pretendeu-se, também, verificar se a exposição resulta em maior grau apenas dos factores temporais e das tipologias disponíveis em determinado momento, ou seja, da estrutura horária da programação ou se, eventualmente, existe alguma maior incidência em determinados conteúdos e canais por segmento considerado nas bases de audimetria.

Estas análises pretenderam a avaliação da existência de comportamentos de “lealdade” a tipologias e canais, assim como a definição de abordagens metodológicas que viabilizem a selecção dos suportes que são necessários para a construção dos planos de *media* em contextos de maior variabilidade dos principais indicadores de audimetria.

Os indicadores em análise correspondem à afinidade, adesão e atractividade de um suporte para um determinado público-alvo e que constam das bases de dados de audimetria.

- O índice de adesão representa a percentagem da totalidade dos indivíduos que contactam com um suporte, que pertencem ao alvo em análise.
- A afinidade entre um alvo e um suporte representa: (1) Indicador do grau de adesão entre o alvo e os contactados por um suporte. (2) Um indicador que compara a percentagem do alvo no suporte com a percentagem do alvo no universo. (3) Um indicador que compara a audiência média de um suporte num dado alvo e no universo.

O Índice de Afinidade compara a percentagem dos contactados por um suporte que pertencem ao alvo (adesão) com a percentagem dos indivíduos do universo que pertencem ao alvo, avaliando em que grau o suporte em causa “atrai” ou “afasta” os indivíduos do alvo em estudo. Se, num dado alvo, a audiência média de um suporte é superior à audiência média do mesmo suporte no universo, então o referido suporte “atrai” mais o alvo em questão do que a globalidade dos indivíduos.

No *Apêndice F.5* pretendeu-se essencialmente avaliar duas questões relacionadas com a selecção dos suportes a incluir nos planos de *media*:

(1) Qual o poder discriminatório das variáveis sociodemográficas ao nível da exposição televisiva a canais e tipologias programáticas? Qual o perfil sociodemográfico de determinados suportes, nomeadamente das tipologias que constam das bases de audimetria e dos canais?

Segundo *Meyer e Hyndman (2005)*, as previsões dos valores de *ratings* devem ser realizadas por segmentos e, posteriormente, os valores médios ponderados.

Foram realizadas análises que avaliam a repercussão dos indicadores sociodemográficos na evolução dos *ratings* por canal e horizonte temporal, que resulta das análises de segmentação das séries temporais. As metodologias simbólicas viabilizam a análise dos índices sociodemográficos associados a objectos de segunda ordem que representam diferentes transmissões de um mesmo programa. A representação da variabilidade associada às diferentes variáveis critério conduz a análises mais relevantes do que considerando simplesmente valores médios observados ou outras medidas de estatística descritiva tradicionais que não permitem uma visão da evolução dos indicadores. Diferentes intervalos de variabilidade de uma variável podem produzir médias equivalentes e pouco representativas da real evolução dos dados.

Verificamos que existe mais-valia em recorrer a estas metodologias no âmbito da segmentação dos padrões de exposição televisiva, assim como para a definição do perfil das audiências de determinados eventos. As abordagens metodológicas apresentadas permitem resultados equivalentes quando os eventos em análise apenas apresentam uma única transmissão.

A caracterização da composição sociodemográfica da audiência da grelha de programas permite a identificação dos espaços comerciais mais adequados à divulgação de determinados produtos ou serviços, tornando as campanhas publicitárias mais eficazes. Permite, também, avaliar e prever o grau de estabilidade que é observado nos *ratings* programáticos através de uma correcta caracterização das audiências de cada um dos canais e posteriormente quanto os comportamentos de lealdade e fidelidade programática. Assim, contribui para a previsão de um dos indicadores que é mais relevantes para este sector. Os *ratings* expressam a dimensão da audiência como uma percentagem do total da população (medição em minutos).

Para a análise dos *ratings*, as metodologias simbólicas pretenderam identificar os segmentos com evoluções mais semelhantes entre si e relativamente ao universo televisivo em função dos canais e das faixas horárias. Atendendo aos mapas de dissemelhança simbólicos, os segmentos que mais se distanciam dos restantes correspondem à classe social AB e à classe social D, às Faixas etárias 4-14, 15-24, 55-64 e +64, à região do Porto e aos Lares com TV Cabo. Os dados para o universo televisivo apresenta uma maior semelhança com os dados dos segmentos adultos e Classe social C2. Através da aplicação de algoritmos de *clustering* simbólicos foi possível identificar para cada canal e faixas horárias quais são os segmentos com uma evolução de *ratings* mais próxima entre si e a construção de macro-segmentos de exposição televisiva, viabilizando a segmentação das audiências televisivas ao nível da evolução dos principais indicadores de audimetria.

Segundo *Webster (1985)* e *Zubayr (1999)*, estas variáveis apresentam capacidade explicativa quando se associam programas a segmentos de telespectadores, contradizendo *Huertas (2002)*, que defende o seu reduzido poder discriminatório dos padrões de consumo televisivo. *Meyer e Muthaly (2008)* salientam que no global verifica-se uma reduzida repercussão ao nível da “lealdade” apesar das diferenças, por vezes, observadas no tempo médio diário de exposição televisiva. Também, referem que as variáveis sociodemográficas apresentam uma reduzida capacidade explicativa da lealdade de curto e longo prazo (coeficientes de determinação no valor de 4,9% e 6,5%), contudo os horários e os canais de maior exposição, explicam uma variação adicional de 18% e 15% da evolução dos indicadores de lealdade.

Considerando a globalidade dos dados existe alguma correlação entres os índices de atractividade e de afinidade associados aos diversos segmentos, mas foi possível realizar algumas associações de canais e tipologias aos segmentos representados nas bases de dados de audimetria. Para programas que apresentam uma só transmissão construíram-se mapas MDS, com posterior aplicação da Regressão Linear (RL) por forma a facilitar a interpretação das dimensões do espaço perceptual. Verificamos que os canais se diferenciam quanto aos programas que maximizam os índices de atractividade.

A análise dos índices de fidelidade, afinidade e adesão, também, permitiu diferenciar os segmentos sociodemográficos relativamente às suas selecções programáticas e ao grau de envolvimento com o meio televisivo e correspondente programação. Para os programas com mais de uma transmissão, uma vez que os indicadores apresentam um intervalo de valores, recorreu-se a metodologias de análise simbólica, nomeadamente às análises factoriais em Componentes Principais. Verificamos que existe alguma heterogeneidade e diferenciação na composição sóciodemográfica da audiência de cada um dos canais, em especial se consideramos os canais públicos versos os canais privados generalistas. As características sociodemográficas em parte permitem diferenciar o público-alvo de cada canal influenciando a evolução e a variabilidade dos diversos indicadores e, por conseguinte, influenciando a posterior previsão dos índices de audiências.

Fernandes (2001, p.116) refere que: “Procurando traçar o perfil do público de televisão, com base nos dados da AGB Portugal (entre 1993 e 1997) por sexo, exceptuando a RTP2 devido ao desporto, são mais as audiências mulheres que vêm televisão nas três estações de televisão”. Para o período em análise verificamos que o Canal 2 minimiza a exposição para ambos os sexos e a RTP1 apresenta uma evolução mais uniforme do decorrer do dia, superando ou igualando os canais privados até às 19H. Nos períodos de maiores audiências os canais privados apresentam uma maior exposição a ambos os sexos, sendo em geral superiores para o sexo feminino. Verificámos com base nas análises às taxas médias de audiências para o total da TV que no final da noite por vezes a exposição do sexo masculino supera a do feminino, tudo será função das tipologias transmitidas.

Observamos um nítido acréscimo nos valores dos *ratings* com o aumento da faixa etária e a diminuição da classe social. A classe social AB não apresenta uma maior exposição no global ao Canal 2, sendo a sua exposição, no decorrer da maior parte do dia, muito semelhante à das restantes classes sociais. Para a classe social AB também se destacam os canais privados, ainda que o diferencial de valores de exposição face ao canal RTP1 não seja tão acentuado. A classe social D apresenta uma elevada exposição a ambos os canais privados, mas muito menor à RTP1.

Todas as classes sociais com a excepção da AB apresentam os maiores níveis de exposição televisiva. Com a redução da classe social maior será o diferencial entre os valores dos *ratings* dos canais públicos versos os canais privados. Os canais privados apresentam maior exposição à classe social D que se diferencia significativamente das restantes. Os resultados do perfil sociodemográfico de cada canal encontram-se condicionados pelas faixas horárias em análise.

Relativamente às faixas etárias, a partir dos 55 anos maximiza-se a exposição televisiva e para todas as idades a exposição é superior aos canais privados generalistas. Para as idades superiores aos 55 anos existe um acréscimo na exposição ao canal RTP1, mas, ficando aquém dos canais privados nos horários pós-laboral.

O Canal 2 não apresenta um grande diferencial de valores entre as diversas faixas etárias. A TVI maximiza a exposição aos mais jovens até aos 14 anos e a partir dos 45 anos e a SIC, também, apresenta maiores índices para os mais jovens até aos 24 anos e a partir dos 55 anos. Segundo *Huertas (2002)*, considerando as faixas etárias o público divide-se em infantil, juvenil e adulto. Verificamos que o padrão de exposição entre os 25 e os 44 anos diferencia-se das maiores faixas etárias.

Para *Fernandes (2001)*, o indicador região geográfica sugere que a RTP1 o Canal 2 são vistos pelo Litoral e Interior; a SIC na Grande Lisboa e Interior; a TVI no Litoral Interior e Grande Lisboa (em termos globais de visionamento é o Litoral e o Interior). Segundo *Huertas (2002)*, ainda que, também, presentes, são menos habituais os trabalhos que comparam o mundo rural e o urbano. O canal SIC para o período em análise apresenta maior exposição aos telespectadores do Litoral Norte e do Interior e a TVI à grande Lisboa e ao Porto.

Em *Huertas (2002)*, defende-se que é necessário incorporar questões que apelam directamente ao modo de recepção para estabelecer diferentes grupos de audiência. Nesse sentido, o critério mais utilizado corresponde à duração do consumo e à distinção entre quem vê muita e pouca televisão, que se apresenta como necessária para estudar e analisar comparativamente assuntos, como, por exemplo, os efeitos.

Em *Fernandes (2001)*, também, constam alguns quadros com a evolução das variáveis ATS e ATV. Na presente tese, foi analisado um outro indicador que corresponde à fidelidade programática. Verificámos que, no canal SIC, os valores são minimizados quando se transmitem filmes, programas matinais, juvenis e programas do final da noite. Com maiores valores de fidelidade destacamos as telenovelas, as séries, os espaços informativos da tarde e os programas adjacentes a ambos os espaços informativos. Para o canal TVI os índices são minimizados por parte dos filmes de fim-de-semana, programas matinais no decorrer da semana, eventos desportivos ou outros, concertos, etc. Maximizam-se para alguns programas juvenis e filmes da noite durante a semana e para programas de muito curta duração.

A RTP1 apresenta maiores valores de fidelidade para os espaços informativos da tarde no decorrer de toda a semana e para os seus programas adjacentes. Os menores valores estão associados aos filmes da noite e da tarde durante toda a semana, programas de debate como o “Prós e Contra” e os programas matinais, entre outros.

Segundo *Fernandes (2001)*, ao nível desta dimensão de construção do conceito de audiência, podem extrair-se duas ilações: a idade e a disponibilidade e predisposição representam indicadores que explicam a prática de visionamento da televisão ao longo do dia. Todavia, há que considerar que, na actualidade, o público é visto como audiência e que, segundo os estudos desenvolvidos nesta área, varia ao longo do dia, ou seja, há uma tendência para uma fragmentação do público de massas. Assim, consoante a audiência disponível em determinado período do dia, a programação incide numa determinada área. Também, foram realizadas análises por faixas horárias e diferenciando a semana do fim-de-semana para verificar quais são os segmentos que em cada período apresentam maiores níveis de exposição e compará-los quanto ao grau de variabilidade observado na exposição televisiva. Para além da idade outras variáveis sociodemográficas condicionam a exposição ao *media* e os padrões de consumo televisivo. Os anunciantes devem considerar os comportamentos de lealdade quando se planifica uma campanha publicitária e foram realizadas associações dos segmentos a programas e tipologias.

A segunda questão em análise apresenta um teor metodológico ao nível da selecção dos suportes: “*Quais as abordagens metodológicas que podem viabilizar a selecção dos suportes a incluir em planos de media?*”

Na opinião de alguns autores, as previsões de *ratings* devem ser realizadas por segmentos e ponderadas para obter estimativas dos valores para o universo televisivo (*Meyer & Hyndman, 2005*):

The results suggest that the best forecast accuracy is obtained from the segment-rating model, used weighted average of these segments forecasts.

This suggests that aggregation bias will be small provided that sufficient demographic data is incorporated via the segmentation. It is recommended that this approach be used to produce the ratings forecasts required by program schedulers and advertisers. ... In order for ratings forecast to be useful to program schedulers the model would to be fitted to recent appropriate data (say last month's data) in order to obtain reasonably accurate forecasts and prediction intervals. (p.16-p.17)

Quanto às abordagens metodológicas que promovam a segmentação das audiências em função dos níveis de exposição foram realizadas análises aos segmentos sociodemográficos que permitiram do ponto de vista das taxas médias para o total da TV e para os *ratings* avaliar quais são os macro-segmentos que apresentam entre si e face ao universo televisivo uma evolução equiparável dos indicadores de audimetria. Se, eventualmente, as previsões dos *ratings* forem realizadas por segmentos estas análises permitem a agregação das observações em função da evolução observada para os indicadores de audimetria. Os *ratings* expressam a dimensão da audiência como uma percentagem do total da população (medição em minutos). Para a análise dos *ratings*, as metodologias simbólicas identificaram os segmentos com evoluções mais semelhantes e face ao universo televisivo, em função dos canais e faixas horárias.

Porém, as análises de segmentação, também, são necessárias em outros contextos e em determinadas etapas necessárias à construção dos planos de *media*. Por exemplo, poderá ser necessário excluir das análises os segmentos que incutem uma maior variabilidade de valores dificultando as análises de previsão e que devem ser previamente identificados. Também, poderá ser necessário considerar estes segmentos por pertencerem ao público-alvo da campanha publicitária, sendo necessário verificar para cada período quais são as sequências de programas que neles predominam, contribuindo, assim, para a selecção dos suportes e adoptando-os para o cálculo dos valores de *ratings*. Os macro-segmentos podem resultar do clustering das sequências identificadas a partir dos ficheiros com a informação base de audimetria e recorrendo a metodologias do âmbito do *Web Mining* e do *Data Mining* Sequencial.

Quando os dados dos ficheiros de base de audimetria não se encontram disponíveis e uma vez identificadas as variáveis de segmentação que apresentam maiores semelhanças (com base em metodologias simbólicas de *clustering* e de dissemelhança) podemos recorrer aos índices de atractividade ou adesão e, assim, identificar quais são as sequências horárias e as sequências programáticas que maximizam a exposição televisiva e que se apresentam como mais prováveis de serem seleccionadas. Para conseguir prever exposições mais concretas em determinada semana os resultados devem ser aprofundados recorrendo a metodologias de *Data Mining* não proposicional e que se descreve no seguinte capítulo.

As metodologias simbólicas viabilizam a análise dos índices sociodemográficos associados a objectos de segunda ordem que representam diferentes transmissões de um mesmo programa. A representação da variabilidade associada às diferentes variáveis critério conduz a análises mais relevantes do que considerando simplesmente valores médios, entre outras medidas de estatística descritiva que não permitem uma visão da evolução dos indicadores. Diferentes intervalos de variabilidade de uma variável podem conduzir a médias equivalentes e pouco representativas da evolução dos dados.

Verificamos que devemos recorrer a estas metodologias no âmbito da segmentação dos padrões de exposição televisiva, assim como para a definição do perfil das audiências de determinados eventos e para a selecção dos suportes a incluir em planos de *media*. As abordagens metodológicas apresentadas permitem resultados equivalentes quando os eventos em análise apenas apresentam uma única transmissão.

Quando for possível recorrer aos dados dos ficheiros de base de audimetria existem abordagens metodológicas mais adequadas à segmentação das audiências face às que constam da bibliografia da especialidade. Nomeadamente as Regras de Associação e as Regras Sequenciais e os respectivos algoritmos de *clustering*, correspondem à melhor perspectiva de análise. Actualmente, a área do *Data Mining* Multirrelacional poderá ainda não ser a melhor opção a considerar em análises de previsão, que envolvem essencialmente valores numéricos, mas a prazo talvez esta situação se venha a inverter.

A caracterização da composição sociodemográfica da audiência da grelha de programas permite a identificação dos espaços comerciais e dos suportes mais adequados à divulgação de determinados produtos ou serviços, tornando as campanhas publicitárias mais eficazes. Permite, também, avaliar e tentar prever o grau de estabilidade observado nos valores dos *ratings* programáticos através de uma correcta caracterização do perfil das audiências de cada um dos canais e quanto aos comportamentos de lealdade e fidelidade programática.

Concluindo, a dinâmica dos comportamentos de consumo televisivo condiciona as abordagens metodológicas de cariz explicativo e de previsão mais adequadas. A selecção poderá ser objecto de uma procura activa, ou não, condicionada pelas características socioculturais e dentro de um padrão de comportamento que procura um determinado conjunto de programas até alcançar determinado equilíbrio. Um indivíduo poderá, também, sentar-se frente ao televisor de forma passiva em função da sua estrutura horária e, não, dos conteúdos disponibilizados.

Será necessário segmentar em faixas horárias a emissão diária em função da variabilidade observada nos principais indicadores de audimetria. Os períodos mais curtos a considerar nas análises posteriores será onde se observa uma maior variabilidade e amplitude de valores. Por vezes, as análises de previsão em cada segmento horário identificado devem ter em conta o comportamento agregado dos telespectadores segundo os seus períodos horários e a sua “lealdade” para com programas em concreto e por possíveis conteúdos dentro dessa tipologia programática. Atendendo às análises de lealdade programática para um segmento de telespectadores existirá uma componente que apresenta maior estabilidade que será ver TV dentro de determinadas faixas horárias e uma segunda componente que poderá ser estável ou mais variável que será ver ou não determinado programa em função da tipologia programática e do seu conteúdo, por exemplo, um filme de determinadas características.

No primeiro contexto será relevante as análises de previsão do âmbito da lealdade programática, recorrendo a métodos de média móveis ponderadas e no segundo contexto provavelmente as Regras de Associação e Temporais permitam alcançar os melhores resultados, considerando, também, os algoritmos probabilísticos e os dinâmicos. Não se deve dar demasiada importância a um programa, mas, sim, ao conjunto de programas com os quais os telespectadores entram em contacto durante determinado horizonte temporal, sendo o *Data Mining* Sequencial relevante neste contexto. O *Data Mining* poderá viabilizar que melhores resultados sejam alcançados quando se pretendem desenvolver modelos de exposição e simultaneamente “optimizando” a selecção dos suportes a incluir em planos de *media*. No seguinte capítulo desenvolve-se uma proposta metodológica que possa acrescentar mais-valias no âmbito do planeamento de *media*.

A duplicação das audiências corresponde a uma área de análise que condiciona o planeamento de *media* e as metodologias de previsão. Sendo o contributo da investigação quantitativa determinante para a resolução de questões que preocupam a vertente empresarial deste sector de actividade, mas, não só, verifica-se, também, uma crescente necessidade de desenvolvimento de análises explicativas e de previsão de audiências recorrendo a abordagens estatísticas (Weber, 2002), que permitam tirar um melhor partido dos dados que quantificam a exposição televisiva, entre os quais se insere o processo de audimetria. As bases de audimetria não devem ser objecto de análise, recorrendo apenas aos métodos mais tradicionais de análise de dados, mas, sim, tendo em atenção as especificidades desta tipologia de dados, tendo-se privilegiado o *Data Mining* e a Análise Simbólica.

V.3.1.4. A interacção entre os comportamentos estruturais

Neste capítulo pretende-se avaliar se existe interacção entre os comportamentos estruturais e responder às seguintes questões:

➤ O público apresenta uma maior passividade ou actividade nas suas selecções? Quais os contextos e comportamentos associados à duplicação primária das audiências? Os *ratings* são influenciados em maior grau por parte da lealdade tipológica, programática ou ao canal? Qual a viabilidade de previsão da exposição recorrendo à duplicação das audiências?

Ao longo das últimas décadas, os índices de duplicação num mesmo canal têm sido atribuídos a inúmeros factores. Podem resultar da lealdade que é dirigida a tipologias e aos canais, ou seja, da tendência do público para apresentar alguma consistência nas suas “preferências” e selecções por determinado canal face aos restantes. Outros autores atribuem-no (*Cooper, 1996*), essencialmente, a fenómenos de lealdade programática e aos efeitos de herança que resultam da passividade dos telespectadores entre programas adjacentes, uma vez que os índices obtidos excedem em dimensão aos que resultariam de uma mera lealdade a canais.

Com base nas análises realizadas os canais generalistas, os canais não se diferenciam significativamente quanto à ordem de grandeza dos índices de duplicação. A evolução dos *ratings*, frequentemente, acompanha a evolução dos índices de duplicação e estes por sua vez estão, maioritariamente, condicionados pelos factores de disponibilidade e de predisposição de exposição ao *media* e em menor grau pelos fluxos de audiências entre canais concorrentes se considerarmos uma grelha já estabelecida.

Por exemplo, o canal SIC em alguns dos anos considerados apresenta maiores valores de exposição aos canais concorrentes e nomeadamente à TVI quando, também, se observam os maiores níveis de TV_OFF. Logo, existe para grande parte do painel uma lealdade ao canal e uma segunda fracção do mesmo, frequentemente, menor, que segue as tipologias que constam do seu repertório num canal concorrente. A lealdade tipológica aparenta ser o factor mais relevante, incutindo nos períodos posteriores alguma passividade, que também decorre da lealdade existente para com o canal, mas os padrões de disponibilidade imperam condicionando fortemente os padrões de consumo televisivo.

Verificamos que a lealdade programática, quantificada através dos índices de duplicação das audiências, não ultrapassa os 50% para o período em análise. Dificilmente, a alteração do horário de transmissão dos programas para horizontes temporais mais tardios consegue reter a audiência, logo os factores temporais imperam anulando uma maior correspondência entre as eventuais preferências e as selecções. A influência das variáveis estruturais é determinante para uma maior compreensão da exposição televisiva e para a respectiva previsão.

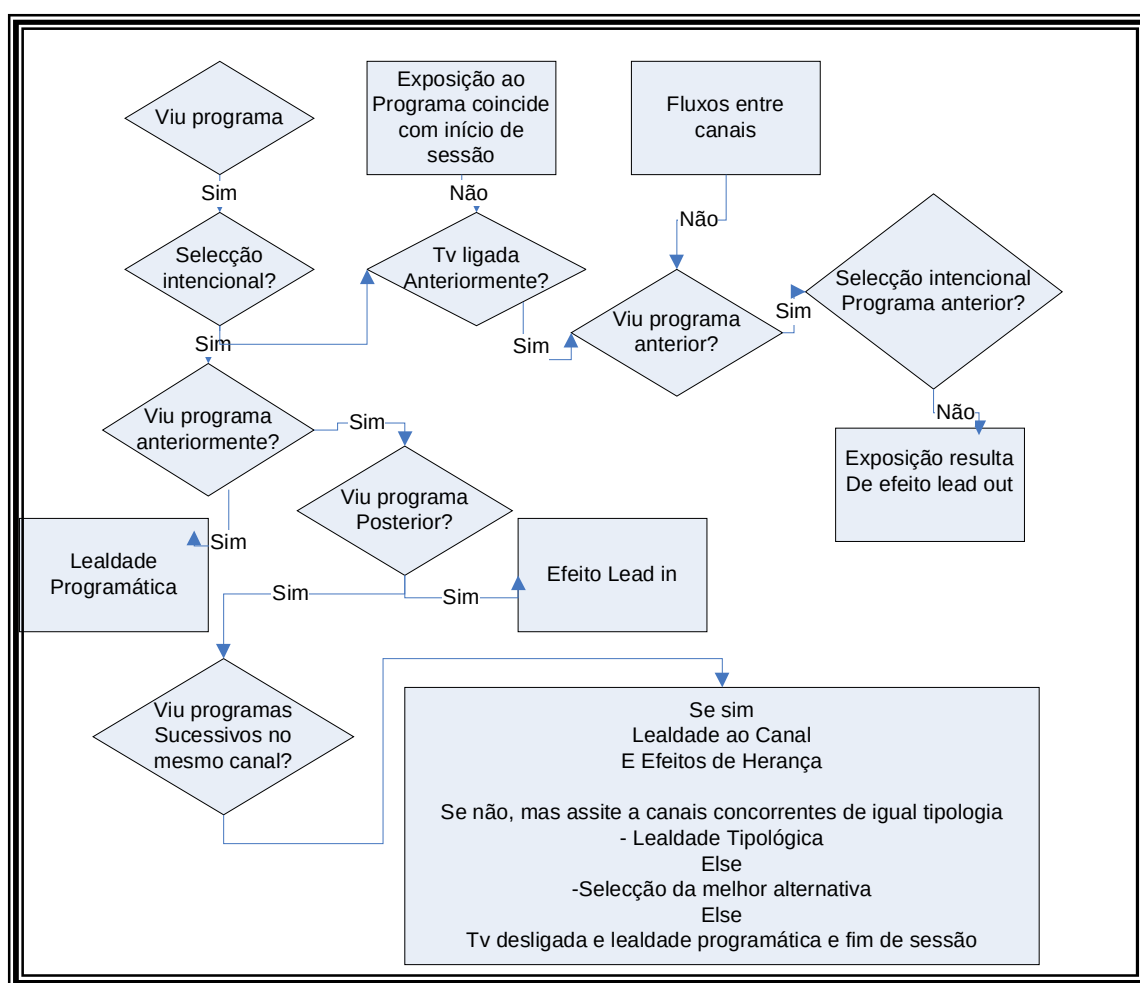


Fig. V.3.1.4.1 – Fluxograma associado à duplicação das audiências.

Fonte: Elaborado pela autora.

Com base nas análises realizadas aos factores estruturais, verificamos que a variável que condiciona em maior grau a duplicação, corresponde ao número de telespectadores que mantém o televisor ligado na interface de dois programas em detrimento dos fluxos para os canais concorrentes.

A nível relativo o grau de duplicação é superior para os maiores valores de TV_ON, ou seja, os períodos que mantêm menor número de telespectadores com o televisor ligado de programa para programa apresentam maior dificuldade em manter o seu público nesse mesmo canal. Em parte do período em análise, para os maiores valores de TV_ON do canal SIC a concorrência será exercida em maior grau por parte de canais por TV Cabo e em alguns dos anos considerados a TVI assume uma maior repercussão.

Para os menores valores de TV_ON será o nível de audiência associado às decisões de desligar o televisor (TV_OFF) que condiciona a duplicação de audiências. Consequentemente, as decisões de ligar e desligar o televisor condicionam em maior grau a dimensão dos *ratings* e os valores de duplicação, em detrimento dos fluxos para outros canais. A duplicação de audiências será superior para menores níveis da taxa média para o total da TV. Os maiores valores dos índices de duplicação primária estão associados aos períodos da manhã, hora do almoço, tarde, seguindo-se o prime time, o início da tarde e noite e o final da noite. O período semanal ou de fim-de-semana não influência, em média, o grau de duplicação de audiências para o período em análise.

Verificamos que existe uma considerável correlação entre as variáveis TV_ON e Dup_SIC, as quais, por sua vez, se correlacionam com os valores de *ratings*. Portanto, a previsão de TV_ON será determinante para a previsão dos valores de *ratings*. As selecções programáticas, para o canal SIC, são, frequentemente, função dos factores de disponibilidade do seu público, verificando-se constatações comparáveis para os restantes canais. As taxas de duplicação encontram-se influenciadas em maior grau pelos factores de disponibilidade, variáveis temporais, fidelidade e atractividade programática.

Na *Figura V.3.1.4.1* e na *Figura V.3.1.4.2* constam os diversos factores que podem influenciar a duplicação de audiências, não sendo a classificação destes comportamentos tão dicotómica como se verifica na reduzida bibliografia internacional da especialidade (*Webster, 2006*).

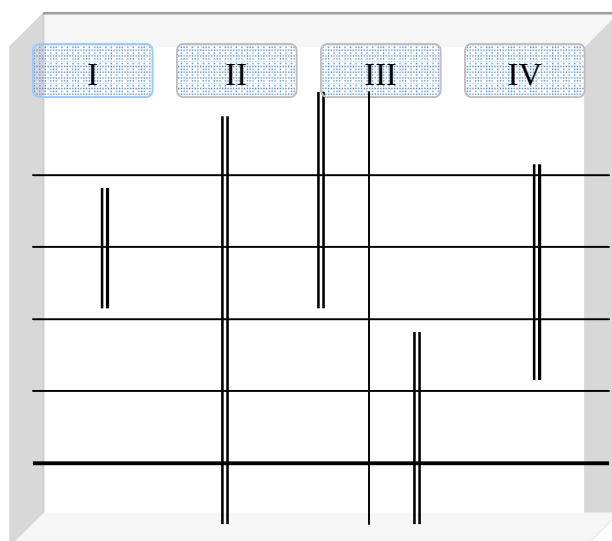


Fig. V.3.1.4.2 – Contextos de duplicação das audiências.
Fonte: Elaborado pela autora.

A duplicação das audiências pode corresponder às seguintes situações :

- (I) Lealdade programática e ao canal;
- (II) Lealdade canal e efeitos de herança;
- (III) Lealdade programática ou tipológica;
- (IV) Lealdade programática e efeitos de herança.

No cenário I apenas se pretende ver determinado programa e no terceiro a duplicação resulta da lealdade a determinadas tipologias e conteúdos, sendo mais irrelevante o canal. No segundo contexto a duplicação decorre da lealdade ao canal, sendo este o nível de decisão de acesso ao *media* e da inércia dos telespectadores. No quarto cenário altera-se o factor que determina o início da sessão televisiva, correspondendo a ver determinado programa.

V.3.2. A apresentação de um modelo estrutural da exposição televisiva

Neste capítulo pretende-se apresentar algumas considerações a respeito da seguinte questão:

- “Em função dos resultados das análises de duplicação de audiências, qual o modelo estrutural que permite caracterizar a exposição televisiva do painel de audimetria? Qual a repercussão ao nível da evolução dos *ratings*; na selecção dos suportes a incluir em planos de *media*; na viabilidade da previsão da exposição televisiva e na selecção dos métodos analíticos posteriormente a adoptar para o planeamento de *media*?”

Podemos identificar na bibliografia da especialidade diversas perspectivas de avaliação dos factores que condicionam a exposição televisiva. Prevalecem os trabalhos associados à perspectiva estrutural e de gratificações e usos do *media*, que se diferenciam significativamente ao nível dos objectivos, fontes de dados, métodos de análise e conclusões apresentadas, nomeadamente no que diz respeito ao grau de actividade subjacente à exposição televisiva.

As duas ópticas, apesar de contraditórias nos seus pressupostos, podem influenciar o consumo televisivo (*Pingree et al., 2001*), defendendo a existência de alguma passividade na decisão de ver televisão, mas, apresentando as selecções, por vezes, uma maior actividade, utilidade e intencionalidade e ser função das preferências por determinados conteúdos. Contudo, uma grande fracção dos trabalhos realizados nesta área salientam que a selecção, também, poderá decorrer de muitos outros factores (*Cooper & Tang, 2009; Cooper, 1996*), assim como as decisões de acesso ao *media*. Do ponto de vista da previsão das audiências, a existência de regularidade do consumo é determinante. Uma maior “passividade” dos telespectadores não é sinónimo de regularidade. Nos contextos de uma maior passividade nas decisões de acesso ao *media*, não se verifica necessariamente, também, uma maior regularidade, uma vez que poderá corresponder a uma forma de preenchimento de um “não tempo” entre actividades de trabalho ou de lazer e que, conseqüentemente, apresenta uma enorme flexibilidade, podendo ser aumentado, reduzido ou anulado no decorrer do tempo.

Uma selecção intencional, também, poderá não estar associada a uma maior regularidade, uma vez que será função da disponibilidade e vontade de ver TV do público. Nas sociedades mais avançadas, verifica-se uma crescente disponibilidade de tempos vazios que correspondem a um entretempo entre o trabalho e o descanso ou entre outras actividades. Conseqüentemente, sendo um tempo restante, não é necessariamente objecto de preferência, desenvolvendo-se esta última, essencialmente, a partir de outras actividades de lazer. Quando alguém assume que em determinado momento apenas o tempo televisivo constitui uma alternativa poderá ser relativamente indiferente a sua materialização.

Neste contexto as preferências ou interesses não são o mais relevante, e o público assistirá aos conteúdos que estão disponíveis em determinado momento independentemente das suas características ou seleccionando-se as “melhores” alternativas disponíveis nesse momento e que podem resultar das inúmeras influências estruturais no consumo televisivo.

Neste âmbito, as influências estruturais podem condicionar fortemente a exposição e a evolução dos dados. Sendo um tempo não preferido, pode ser substituído com muita facilidade, inclusive de forma não consciente, inculcando irregularidade à exposição televisiva. Para determinados segmentos com uma maior disponibilidade e com uma vida quotidiana caracterizada em maior grau por determinadas rotinas diárias ou semanais, independentemente de exercerem uma actividade laboral ou não poder-se-á verificar o oposto, existindo uma maior regularidade de tempos vazios no decorrer do tempo.

O tempo televisivo pode ser adaptado aos ciclos laborais, ganhando de forma crescente em duração quando estes se modificam e aos fins-de-semana ou durante os períodos de férias pode ficar facilmente substituído pelo tempo de lazer ou tempo livre. Quanto maior for o tempo livre menor será o consumo televisivo e tendencialmente poderá ser mais irregular. Verificamos, por análise das taxas médias de audiências para o total da TV, que nos períodos de maiores audiências a exposição global ao *media* apresenta uma maior estabilidade ao fim de semana, mas em outros períodos horários verifica-se o inverso.

Logo, as faixas horárias da tarde e da manhã destinam-se a outras actividades de lazer, inculcando irregularidade à exposição e aos indicadores de audimetria que dela resultam. O tempo vazio experimenta-se repetidamente sempre que se repete o tempo quotidiano ligado, essencialmente, ao ciclo laboral, mas poderá não apresentar regularidade. Quando o tempo de contexto é o tempo de lazer, o tempo televisivo encontra maiores dificuldades em se conseguir implantar, sendo a exposição global ao fim de semana mais variável no decorrer do período matinal e em determinados períodos da tarde.

Verificou-se uma reduzida duplicação das audiências entre os suportes transmitidos no decorrer da tarde de fim-de-semana para fim-de-semana.

O tempo vazio é característico em maior grau dos estratos sociais menos favorecidos. Quem tem menos possibilidades ao nível do capital económico, social e cultural convertem o seu tempo num tempo vazio e que não tem valor. A classe social D que apresenta a menor capacidade económica e nível de instrução, segundo as análises realizadas, apresenta os maiores níveis de exposição ao *media*, assim como o sexo feminino e as faixas etárias a partir dos 55 anos. Contudo, ao nível da variabilidade da exposição não se verifica uma correspondência entre os maiores níveis de exposição e uma menor estabilidade dos valores, se considerarmos equivalentes períodos.

Se compararmos a semana e o fim-de-semana verificamos que os segmentos com uma menor exposição, também, apresentam uma exposição mais uniforme no decorrer de toda a semana. O tempo vazio flutua no decorrer da semana e opõe-se ao tempo livre e como tempo não preferido não implica a necessidade de as selecções serem realizadas em função dos conteúdos, especialmente ao nível das maiores faixas etárias, enquanto os mais jovens tendencialmente procuram tempos televisivos com conteúdo e um consumo mais intensivo. Quando se vê televisão no decorrer do tempo de descanso, de não fazer nada, alguns segmentos acabam por lhe exigir muito pouco e será um tempo de entretenimento que completa a espera entre as obrigações, sendo a ficção o género televisivo que maximiza a exposição para o período em análise.

Frequentemente, a relação para com o *media* não passa pela atractividade exercida por determinados conteúdos e pelas tipologias programáticas. Apenas se pode reconhecer determinadas tendências de “gostos” de determinados segmentos em relação aos restantes considerando alguns escassos géneros televisivos, como, por exemplo, a associação entre os eventos desportivos e o segmento masculino e que se constatou nas análises factoriais e de MDS simbólicas onde se pretende avaliar a exposição por segmento sociodemográfico representado nas bases de dados de audimetria.

Constatou-se alguma dificuldade na associação de conteúdos às diversas variáveis de segmentação das bases de dados de audimetria que correspondam a classificações realmente discriminatórias dos padrões de consumo televisivo. Podemos, sem dúvida, identificar inúmeras perspectivas explicativas da exposição ao *media*, ao nível do grau de actividade versus passividade nas decisões de acesso ao *media* e na selecção dos conteúdos.

Segundo *Cooper e Tang (2009)*, a exposição é essencialmente instrumental, ou seja, associada a hábitos de consumo televisivo, o que pressupõe alguma regularidade subjacente e salientam que a audiência poderá ser classificada como activa em “ambientes estruturais”, devendo-se abandonar a dicotomia de passividade ou/versos de actividade. Porém, certamente que para determinados segmentos de telespectadores a selecção também poderá decorrer de hábitos desenvolvidos e condicionados pelas estruturas existentes, incidindo em determinado canal, conjunto de canais, tipologias de programas ou programas.

A perspectiva estrutural apresenta alguns factores condicionantes da exposição televisiva que podem contribuir de forma mais significativa para a previsão da exposição (*Cooper, 1996*). No segmento feminino e em especial de estratos menos favorecidos a existência do tempo vazio e uma preferência pelo género narrativo, como as telenovelas e os filmes e os “reality shows”, que dramatizam aspectos da vida real, poderá incutir uma maior regularidade na exposição quando estas tipologias são transmitidas.

As análises de lealdade programática permitiram avaliar a duplicação das audiências entre transmissões de um mesmo programa ou espaço televisivo, sendo as telenovelas a tipologia de programas que maximiza estes indicadores, mas que, no entanto, ficam aquém de metade dos telespectadores que assistem a determinado episódio e os *ratings* apresentam valores que no geral são inferiores aos 30%.

Os espaços de transmissão de filmes apresentam reduzidos valores de duplicação e em geral inferiores a 25%, o que permite confirmar a existência de irregularidade na exposição, independentemente dos factores e das perspectivas explicativas que a condicionam maioritariamente. Os espaços informativos, também, apresentam valores de duplicação inferiores aos 40%, apesar de lhe serem associadas as motivações instrumentais, ou seja, a procura de conteúdos a esta tipologia de programas.

A *Figura V.3.2.1* representa os diversos contextos de exposição intencional ou não intencional e de ausência de exposição que se podem verificar no painel de audimetria, onde fica bem patente que a exposição em inúmeros contextos não decorre certamente das preferências dos telespectadores e existe uma maior prevalência das influências estruturais, como, por exemplo, dos factores de disponibilidade, o acesso e o conhecimento ou inclusive do grupo em que o telespectador se encontra inserido e das práticas quotidianas de cada indivíduo. A mesma Figura ilustra e resume os diferentes contextos de exposição que podem ocorrer e que requerem diferenciadas perspectivas de análise quando se pretende tentar prever e explicar a exposição. A exposição poderá ser intencional ou não, assim como quando se evitam determinados conteúdos.

Não existe consenso na bibliografia da especialidade relativamente ao grau de actividade e intencionalidade subjacente às decisões do que ver. Inclusive um mesmo indivíduo pode no decorrer do tempo e em função dos contextos apresentar uma maior ou menor actividade e regularidade de comportamentos de consumo televisivo. As variáveis sociodemográficas podem não representar na realidade um perfil realmente discriminatório dos padrões de consumo televisivo, devendo ser complementadas com variáveis psicográficas.

Na presente tese, defende-se que a audimetria permite identificar da forma mais objectiva possível as selecções realizadas, e que podem resultar dos seguintes factores: de uma real preferência ou interesse por determinado programa ou tipologia; na escolha da alternativa que é considerada como sendo a melhor face aos conteúdos programáticos disponíveis em determinado momento, sendo a TV consumida como um produto em si, que visa preencher os tempos de lazer, o tempo vazio ou como forma de evitar o isolamento e solidão, entre outros muitos factores.

Alguns telespectadores em primeiro lugar podem decidir ver ou não TV e, posteriormente, o que ver. Porém, também, existem outros níveis de decisão, como, por exemplo, a selecção de uma tipologia, programa e fracção de programa que pode decorrer de comportamentos estruturais ou de um “interesse” e uma “preferência”. Por vezes, a selecção poderá ser realizada sem ter em atenção os conteúdos e ser função das influências estruturais, uma vez que ver TV corresponde a uma experiência de preenchimento de um tempo de descanso, de convívio familiar, de falta de ocupação, entre horas de estudo e trabalho ou de actividades domésticas, existindo algum distanciamento face ao *media*. A posição face ao “não tempo” pode marcar diversos estilos e formas de consumo.

Diversas estratégias de programação são utilizadas para tentar reter o público por um maior horizonte temporal no decorrer do tempo de lazer ou vazio e a exposição poderá resultar das influências estruturais decorrente dos factores de planeamento, por exemplo, das estratégias de “Lead_In” e de “Lead_Out”, entre outras. A estruturação da grelha condiciona a evolução dos fluxos de audiências entre programas. A falta de um real interesse por parte dos conteúdos transmitidos torna a exposição televisiva fortemente dependente dos factores de disponibilidade e da “vontade”, incutindo maior ou menor irregularidade no consumo televisivo a nível individual e agregado e, consequentemente, dificultando fortemente as análises de previsão da exposição individual, agregada e dos indicadores que dela resultam.

Verificamos que as taxas médias de audiências apresentam uma maior estabilidade quando estão associadas a faixas horárias, e a incapacidade das estações televisivas para conseguir manter os níveis de audiências quando os programas são alterados na grelha. Por conseguinte, a exposição televisiva encontra-se fortemente condicionada pelos factores de disponibilidade, pela estrutura horária dos programas, entre outros factores estruturais e assim como se verifica a nível internacional devem certamente imperar as motivações ritualísticas para a maioria das tipologias programáticas transmitidas, às quais não podemos necessariamente associar uma maior regularidade da exposição televisiva.

Do ponto de vista da previsão, como se demonstrou, anteriormente, as análises não devem ser realizadas com base nos interesses, intenções e inclusive comportamentos passados caracterizados por frequências de exposição e baseadas em análises probabilísticas que associam o conceito de probabilidade a frequência de exposição (Zarate, 1997). Uma vez que a selecção e o acesso ao *media* em alguns contextos poderá não ser intencional, este processo sublinha a debilidade de considerar as preferências e as motivações como o único componente da caracterização do consumo. Nesse sentido, as análises enquadradas na teoria de gratificações e usos do *media* são, por si só, insuficientes, não só para a caracterização do consumo televisivo, como, também, totalmente inadequadas, por si só, para a previsão.

A exposição no global, também, pode estar associada a uma maior passividade dos telespectadores e às influencias estruturais. Nos contextos em que, eventualmente, exista uma maior actividade nas selecções será necessário avaliar a sua repercussão ao nível da evolução dos indicadores de audimetria e nas abordagens analíticas, que possam promover a sua posterior previsão. A análise de fluxos de audiências entre os canais viabiliza a sua avaliação e consta das análises realizadas. Podemos, sem dúvida, identificar formas de consumo mais estáticas e outras dinâmicas, por exemplo, onde incidem os comportamentos de “zapping”.

O “zapping” associa-se, essencialmente, aos mais jovens e aos estratos sociais mais elevados, enquanto críticos da TV actual. Este comportamento caracteriza-se pela contínua mudança de canal, na procura da mudança e é praticada por quem identifica facilmente a estrutura da programação, relacionando-se de forma simultânea com vários suportes, e deste modo alcançam o objectivo da máxima satisfação a partir da experiência televisiva.

Estes telespectadores relacionam-se em maior grau com os conteúdos e rejeitam o tempo vazio, consumindo a TV de forma dinâmica e muito mais intensa. Os espaços comerciais convertem-se em obstáculos e verificámos de acordo com as análises realizadas que o decréscimo observado nas audiências das inserções publicitárias pode ser significativo, sendo especialmente difícil a identificação dos factores que para tal contribuem.

Uma vez que a exposição não é, frequentemente, intencional e quando o é, e resulta da procura de determinados conteúdos dificilmente será conhecida algumas semanas antes de ocorrer, as sondagens de intenção de ver produzem inevitavelmente resultados muito insatisfatórios. Por conseguinte, é necessário recorrer a análises matemáticas que permitam alcançar os melhores resultados em função do grau de regularidade observado nos índices de audiências e em alternativa considerando os dados dos ficheiros de base de audimetria, ainda que esta perspectiva de análise seja muito mais demorada e, consequentemente, menos prática para quem precisa realizar previsões a curto prazo e sistematicamente.

Outro aspecto a ter em consideração, resulta de não existir qualquer consenso relativamente à influência das variáveis sociodemográficas na exposição televisiva, tendo sido avaliada a sua repercussão na evolução dos “*ratings*”, nomeadamente nas taxas médias de audiências e nas taxas médias de audiências para o total da TV. Alguns autores salientam o papel da TV como meio de entretenimento, que abrange os mais instruídos, assim como os menos escolarizados, e outros sectores da população, ou seja, uma uniformização das motivações subjacentes à exposição televisiva, mas que,

contudo poderá ser caracterizada através de diversas formas de consumo e, por vezes, contraditórias. Será expectável que as principais motivações correspondam ao preenchimento de tempos de lazer e de tempos vazios, de acordo com a perspectiva estrutural, assim como à procura de informação e de outros conteúdos e numa forma de fuga à solidão.

A audimetria permite que os padrões de exposição televisiva sejam, caracterizados com um maior rigor e detalhe, sendo necessário recorrer a outras fontes de dados quando se procuram motivações, razões e consequências que derivam da sua ocorrência. A audimetria viabilizou o desenvolvimento da investigação quantitativa das audiências que apresenta como principal objectivo a quantificação das audiências essencial para o planeamento de *media*. No texto apresentado em *Fernandes (2001)* consta que:

O discurso dos ratings refere-se aos programas vistos, mas o que poderá (previsão) vir a ser um programa de sucesso, não é controlado, pois existe a crença de que o comportamento das audiências se mantém sem grande variação. O telecomando veio alterar este factor “inércia do telespectador”, diminuindo o controlo da televisão sobre o comportamento do público. (p.104)

Com base em toda a exposição realizada relativamente aos modelos probabilísticos frequentistas de previsão das audiências que constam da bibliografia da especialidade (*Zarate, 1997*), ficou demonstrado o grau de incorrecção que existe nesta afirmação, sendo esse o pressuposto da investigação das audiências pré-audimetria e que recorre aos inquéritos que apenas permitem uma caracterização pontual e subjectiva da exposição televisiva. A investigação de audiências realizada antes do desenvolvimento dos audímetros e cujo método de recolha de informação são os inquéritos parte do pressuposto de regularidade dos comportamentos e dos conteúdos, o que pode carecer, actualmente, de consistência, como também, constatamos nas análises realizadas no âmbito da duplicação das audiências.

Verificámos que uma percentagem inferior a metade do painel dos que assistem a determinado programa poderá apresentar uma maior regularidade na exposição, dia após dia, denotando a possível existência de alguns hábitos de consumo televisivo.

Uma vez que no geral os valores dos *ratings* apresentam valores moderados conclui-se que a irregularidade impera maioritariamente nos padrões de consumo televisivo do painel de audimetria. Inclusive os espaços informativos apresentam moderadas taxas de duplicação de audiências, contrariando as afirmações de que nesta tipologia de dados incide a exposição de segmentos que apresentam um reduzido consumo televisivo, com uma actividade quotidiana predominantemente extra-doméstica e que estes conteúdos representam o que existe de cariz mais ritual e regular no processo quotidiano.

A investigação tradicional aceita a hipótese de regularidade, uma vez que não existem recursos, quer a nível económico, quer tecnológico que viabilize estudar a audiência no decorrer de longos horizontes temporais. Assim, nesta tese privilegiou-se uma fonte de dados que viabiliza uma análise diacrónica das audiências da forma mais objectiva e abrangente.

Para o segmento que apresenta uma exposição mais regular, as selecções podem ser mais consistentes ou não. Os telespectadores podem sistematicamente seleccionar o mesmo programa ou variar as selecções no decorrer do tempo, escolhendo o que consideram mais apelativo ou o que menos lhe desagrada face à oferta disponível ou sem ter em atenção os conteúdos, prevalecendo por si só a “experiência televisiva”. Aparentemente, atendendo aos índices de lealdade programática quando existe uma maior regularidade na exposição também implica uma maior estabilidade nas selecções programáticas realizadas.

A prática televisiva transforma-se na prática presente quando não existem outras alternativas, estando directamente relacionada com o tempo e produzindo-se um distanciamento face aos conteúdos televisivos. A TV poderá ser um dispositivo rutinizador em especial para quem tem uma vida laboral onde não prevalece a rotina. Segundo as análises realizadas às taxas médias de audiência para o total da TV, a classe social AB que se caracteriza por uma maior capacidade económica e grau de instrução apresenta os menores valores e também uma menor variabilidade na exposição televisiva no decorrer de toda a semana.

A classe social D apresenta os maiores valores de exposição televisiva, subindo significativamente ao fim de semana. Os conteúdos, apesar de manterem a TV em funcionamento, podem não ser o objectivo central da relação do público com o televisor. Porém, quem apresenta maior irregularidade na exposição deverá ser também muito mais variável nas escolhas realizadas no decorrer do tempo. Por exemplo, os jovens caracterizam-se por um perfil mais intensivo de consumo no sentido de “assistir aos conteúdos” e, não ao *media*, ou seja, relaciona-se com a TV a partir da programação, mas com uma reduzida fidelidade perante programas específicos, constatada nas análises realizadas ao indicador que quantifica a permanência das audiências no decorrer dos programas.

Segundo *Copper (1997)* e *Gensh e Shaman (1980)*, a exposição televisiva apenas depende dos conteúdos disponíveis, e, não, de numa necessidade activa dos telespectadores. *Webster e Wakshlag (1983)* referem que os factores estruturais podem viabilizar a previsão da exposição. O planeamento e a disponibilidade do público determinam as selecções e a estrutura das opções disponíveis envolvendo as diversas tipologias de comportamentos de duplicação das audiências poderá permitir prever a exposição.

Também, verificamos que impera a influência da exposição global ao *media*, a qual coloca grandes dificuldades quando se pretende promover a sua previsão. A irregularidade da exposição televisiva condiciona e atenua a capacidade predictiva das restantes variáveis estruturais, existindo sem dúvida uma significativa influência das variáveis de planeamento.

V.4. A VIABILIDADE DE CONSTRUÇÃO DOS MODELOS DE EXPOSIÇÃO RECORRENDO AOS MODELOS DE “RATINGS”

No passado, para construir um plano de *media* tem sido privilegiada a previsão dos valores de *ratings*. Em alternativa será necessário a selecção de um conjunto de suportes publicitários que permitam determinada cobertura e frequência de exposição, em função das restrições de cariz económico, mas, não só, o que implica a necessidade de

desenvolver modelos de exposição ao *media* recorrendo às bases de dados de audimetria.

Não será viável a construção de planos de investimento em publicidade televisiva recorrendo aos inquéritos e às análises probabilísticas de previsão segundo o conceito de frequência de exposição (Zarate, 1997). Também, não será viável a construção de uma base de dados que caracterize o comportamento individual e que possa, simultaneamente, apresentar alguma precisão considerando a maior parte do painel de audimetria.

Os inquéritos caracterizam o consumo televisivo de forma pontual e baseiam-se no pressuposto de que o período seleccionado é representativo, a nível médio, do consumo televisivo anual ou no mínimo referente a um período relativamente alargado, e de acordo com a sazonalidade que caracteriza o consumo televisivo. Nas análises realizadas antes ao desenvolvimento da audimetria pressupõe-se que existem hábitos de ver TV e que estes não se alteram significativamente, ou seja, aceita-se a existência de uma significativa regularidade da exposição ao *media* e das selecções realizadas.

A audimetria apresenta a grande vantagem de permitir uma caracterização detalhada da exposição no decorrer de um longo horizonte temporal. Não é relevante conhecer os dados apenas para um breve período quando os hábitos tendem a alterar-se, ou não existem, verificando-se significativas diferenças entre a audiência de cada semana, e que as análises de duplicação das audiências que foram realizadas permitiram confirmar.

A investigação mais tradicional que recorre aos inquéritos aceita que os comportamentos são regulares e que na audiência de cada programa sempre constará um número aproximado de indivíduos com um semelhante perfil sociodemográfico. Consideram nessa audiência, a percentagem do público que assiste a determinado programa com diferentes frequências de exposição no decorrer do tempo.

As análises realizadas aos indicadores acumulados de audimetria permitiram demonstrar que a audiência é, frequentemente, muito variável para uma elevada fracção do painel, e logo invalida as perspectivas tradicionais realizadas com os inquéritos e baseadas nos princípios de regularidade dos comportamentos. A regularidade não se cumpre devido aos diversos factores estruturais e individuais que condicionam a exposição televisiva, alguns dos quais foram avaliados na questão anterior.

A audimetria permite ao sector televisivo dispor de um instrumento tecnológico que viabiliza uma grande diversidade de análises, e que caracteriza determinados comportamentos que dificilmente poderiam ser avaliados recorrendo a outras fontes de dados, como, por exemplo, o “zapping” .

Como qualquer outra fonte de dados apresenta as suas vantagens e desvantagens. Porém, corresponde ao único método que permite uma análise da evolução dos padrões de consumo televisivo. Na tecnologia de audimetria de medição das audiências existem mecanismos de controle de qualidade, como, por exemplo, os controles externos, onde todos os dias são distribuídas informações detalhadas aos seus membros relativamente à produção dos dados e gestão dos lares que integram o painel. Para avaliar a fiabilidade dos dados recorre-se a métodos alternativos de medição das audiências, como, por exemplo, aos estudos coincidentais e aos diários de audiências.

Verificou-se uma evolução ao nível das metodologias de construção dos modelos de exposição que quantificam as audiências. Quando os dados de audimetria começaram a ser disponibilizados, inicialmente era necessário recorrer a fórmulas adhoc de duplicação que permitissem a quantificação da distribuição de frequências. Posteriormente, com dados de exposição a nível individual, os algoritmos de optimização passaram a incidir nos dados do mês anterior, todavia esta abordagem apresenta algumas das limitações apresentadas pelos inquéritos, ou seja, o pressuposto de estabilidade dos comportamentos.

Por exemplo, em *Craig e Ghosh (1993)*, consta a apresentação de um modelo de optimização combinatória que representa uma função objectivo que pretende maximizar o número de indivíduos que são expostos ao plano de forma “efectiva”. Dada uma matriz de caracterização da exposição histórica e determinadas restrições orçamentais pretende-se seleccionar a combinação de suportes disponíveis que permitem maximizar o número de telespectadores que lhe são expostos pelo menos m vezes e que são ponderados em função da semelhança das suas características face ao público-alvo seleccionado.

Os modelos dos efeitos da exposição publicitária são relevantes na medida em que quando se selecciona um plano deverá ser atribuída uma maior ponderação aos lares ou indivíduos que comprem a marca acima da taxa média observada. Apesar de alguns autores defenderem que o número de exposições deverá ser no mínimo de 3 (*Krugman, 1972*), também, existem outras perspectivas que defendem que o seu valor deverá ser uma função dos factores de Marketing e de Comunicação, como, por exemplo, ser uma marca nova ou estabelecida, o grau de lealdade à marca e a duração da campanha (*Ostrow, 1982*). As marcas com uma maior quota de mercado requerem um menor número de exposições e a natureza do mercado alvo, também, influênciam o valor a adoptar. A complexidade da mensagem publicitária, a sua duração e as características que a identificam, assim como o produto e serviço a divulgar, também, influenciam o valor do número de exposições que deve ser considerado como o mais adequado.

Para além das restrições orçamentais, também, podem ser definidas restrições quanto ao número de inserções a considerar em determinados canais, faixas horárias ou tipologias programáticas e que são função das análises realizadas ao perfil sociodemográfico dos suportes e dos comportamentos de lealdade. Tradicionalmente, tem-se utilizado para a quantificação da exposição a curto prazo, as bases de dados que representam a exposição do mês anterior (*Craig & Ghosh, 1993*), onde, por vezes, é necessário recorrer à distribuição Beta Binomial Negativa para tentar calcular a informação em falta nas bases de dados para alguns dos telespectadores.

Porém, uma vez que não existe estabilidade e regularidade na exposição televisiva esta abordagem não é, actualmente, válida. Atendendo às dificuldades subjacentes à construção de uma base de dados de exposição que represente as selecções futuras de cada telespectador, a construção de um plano passa a implicar a necessidade de calcular os valores de *ratings* de cada inserção a contratar e dos suportes em que estas estão inseridas, uma vez que o seu somatório nos proporciona uma medida do valor das audiências acumuladas.

Em alternativa, seria necessário recorrer a determinados métodos de probabilização dos futuros comportamentos, mas que apresentam na opinião de *Monkman (1996)*, grandes limitações, entre as quais:

The answer given by any such probability model will inevitably be the average performance of all schedules meeting the definition specified. In fact there is still plenty of scope within a channel/ daypart allocation for an individual schedules` performance to vary significantly from the average. (p.2)

Uma das alternativas adoptada em maior grau pelo sector será tentar prever os valores de *ratings* recorrendo aos dados históricos. Consequentemente, também, são necessárias abordagens que permitam calcular os valores de *ratings* com as menores margens de erro possíveis ou eventualmente encontrar alternativas de análise que se descrevem na seguinte questão em avaliação.

Uma vez que os suportes publicitários podem ser adquiridos de forma isolada no “spot market”, sendo nestes contextos essencial a previsão dos valores de *ratings* também se deve considerar esta perspectiva de análise, nomeadamente as análises univariadas e multivariadas de previsão e explicativas dos principais indicadores de audimetria, entre os quais se destacam os *ratings* e as taxas médias de audiências para o total da TV ou os *shares* de audiências. Os algoritmos de previsão podem ser, essencialmente, agregados em dois grandes conjuntos, por um lado nas análises matemáticas, e por outro nas sondagens de intenção de ver, entre outras abordagens qualitativas.

As análises matemáticas podem estar baseadas na extrapolação de séries cronológicas, ou ter em conta variáveis explicativas exógenas, como, por exemplo, as características dos conteúdos transmitidos e das situações de concorrência. Nos modelos causais impera a aplicação das metodologias de Regressão Linear, tendo-se nesta tese abordado as eventuais mais-valias da aplicação das metodologias Simbólicas. Também, podemos recorrer a técnicas de Inteligência Artificial, como, por exemplo, às Redes Neurais ou às Árvores de Regressão que também foram consideradas, uma vez que não se pretendeu apenas “prever” audiências, mas, sim, procurou-se perceber o que condiciona a exposição e a evolução dos indicadores, recorrendo a análises não lineares que também permitem avaliar as interdependências existentes entre os diversos factores.

Os métodos qualitativos, como, por exemplo, o “Brainstorming”, “Delphi”, a construção de cenários históricos ou o “Cross-impact” utilizam-se nas situações em que o passado não nos proporciona uma informação directa sobre o fenómeno em avaliação, como, por exemplo, para a análise das audiências de um novo programa, canal ou tipologia programática. Estes métodos são mais úteis em situações em que os dados são escassos, ou em que, pela incerteza e irregularidade dos dados e do ambiente, a extrapolação assume uma validade muito duvidosa. O método *delphi* corresponde a um processo iterativo de convergência de opiniões e o método da analogia histórica pode ser utilizado quando é conhecida a evolução que decorre da introdução de produtos ou serviços análogos e é utilizado no campo da audimetria.

Os sistemas qualitativos de sondagens de intenção de ver, frequentemente, apresentam grandes margens de erro, uma vez que a exposição poderá não estar associada a uma intencionalidade e nas estações televisivas verifica-se, por vezes, uma tendência para ocultar a programação, para além dos problemas logísticos associados a este processo e a falta de capitalização do conhecimento que o convertem num dos métodos mais onerosos. Este método é, essencialmente, utilizado nos períodos de maiores audiências e mais irregulares, mas, não, produzem bons resultados, uma vez que os telespectadores acabam por manifestar os seus interesses, mas a exposição também se encontra influenciada por variáveis estruturais.

As previsões qualitativas, recorrendo à intuição e a juízos de valor, também devem ser apenas um complemento e apresentam uma maior relevância quando se pretende a análise de novas grelhas de programação e a sua posterior gestão e em menor grau para a actividade de planeamento de *media*, onde é necessário estimar uma grande diversidade de indicadores e recorrer à análise de bases de dados de grande dimensão. Quando se constroem planos recorre-se a esta abordagem, essencialmente, quando vão ser transmitidos programas que podem gerar observações atípicas, sendo necessário dispôr de informação sobre a audiência destes programas e da programação concorrente para se poder realizar uma boa negociação com a estação televisiva.

Para além dos dados históricos, também, se deverá dispor de informação “qualitativa” sobre a sua mais provável repercussão na evolução dos dados, para assim se justificar um menor investimento na compra dos suportes publicitários associados a suportes mais “atípicos”, como, por exemplo, determinados eventos desportivos. Determinados eventos podem incutir um grande acréscimo nos níveis de exposição e que são conseguidos à custa dos telespectadores dos canais concorrentes e da adesão de público que não assiste com regularidade ao *media*. Sem se dispor de alguma informação não sabemos qual será o melhor investimento a realizar.

As abordagens quantitativas de previsão são inúmeras, desde as mais simples extrapolações ou prolongações de tendências, até à análise da grelha de programação e à identificação de variáveis explicativas com capacidade predictiva. Podemos utilizar análises matemáticas que inter-relacionem e ponderem factores previamente identificados, assim como recorrer a sistemas de simulação que permitam identificar regularidades estatísticas. A selecção do método de previsão mais adequado não depende apenas da precisão das análises (*Tavares et al., 1996*):

A selecção do método mais adequado, está longe de se esgotar em meras questões de precisão dos resultados, ... adicionalmente, questões como a disponibilidade de dados, a facilidade de uso e implementação, os custos envolvidos (de desenvolvimento, parametrização e implementação dos modelos, de recolha e manutenção dos dados necessários, etc.), restrições de tempo e prazos, jogam um papel fundamental na selecção do método mais adequado, apreciação que depende marcadamente do contexto. (p.266)

Consequentemente, a previsão de *ratings* ou da exposição televisiva poderá ser realizada através de diversas abordagens que na realidade devem ser combinadas, já que nenhuma permitirá que sejam alcançados resultados com elevada precisão. Devemos ter sempre presente que as previsões deterministas não são viáveis, ao contrário, do que se verifica, frequentemente, nas Ciências Exactas.

Of course advertisers do not expect the predicted GRP's to be exactly the same as the eventual GRPs, and moreover, both are subject to sampling error, as they are derived from a panel. Hence some tolerance is permitted, with typical thresholds for the under-delivery in total campaigns *ratings* being 3-5%. Also note the GRP comparisons are made for the entire campaign, not on a program-by-program basis. (Danaher , 2011, p.21)

As abordagens quantitativas partem do pressuposto de que se dispõe de informação passada a respeito do fenómeno em estudo. Facilmente, identificamos séries temporais onde os valores passados são de grande utilidade, por existir um padrão de comportamento fixo, enquanto que outros são meramente aleatórios. Nas séries económicas identificamos uma componente determinista e outra aleatória.

Os métodos quantitativos podem ser univariados, como, por exemplo, os métodos de decomposição, os modelos de amortecimento exponencial e os modelos ARIMA univariados ou podem ser “causais” em que a explicação advém de considerar factores externos não controláveis ou internos e as relações identificadas não são obrigatoriamente relações de uma real “causalidade”, mas que se devem manter no futuro. Os factores a considerar nas análises multivariadas devem ser mais fáceis de prever do que a variável dependente ou serem conhecidos. Nos modelos causais destacam-se as análises de Regressão Linear, os modelos de Regressão Dinâmica e os modelos de Transferência que permitem analisar a relação dinâmica entre as séries temporais, assim como os métodos Bayesianos. Os métodos univariados estritamente extrapolativos limita a sua utilização directa a ambientes mais estáveis, mas podem ser utilizados para a previsão de factores a utilizar nos métodos causais e no estabelecimento de tendências em análises de sensibilidade.

Quando se pretende a análise dos indicadores de audimetria coloca-se a questão de quais são as metodologias a adoptar, as quais são função dos contextos em análise e já identificados nas análises descritivas e explicativas realizadas à evolução dos dados. De acordo com as análises realizadas, facilmente, identificamos uma grande diversidade de contextos de previsão onde encontramos situações em que se observa uma reduzida variabilidade nos dados, e que consequentemente são mais, facilmente, previsíveis, assim como, situações onde a variabilidade é muito superior, e que nos colocam maiores dificuldades. Por exemplo, as faixas horárias e o programa a prever, também, influenciam a abordagem a ser seleccionada, nomeadamente quando nos debruçamos sobre os eventos desportivos, que correspondem a uma das tipologias de programas onde as características dos conteúdos transmitidos podem influenciar em maior grau a exposição televisiva e a evolução dos *ratings* de determinado canal e, por vezes dos canais concorrentes.

Consequentemente, iniciaram-se as análises com a caracterização da variabilidade dos principais indicadores de audimetria quando estão associados a programas ou a faixas horárias, uma vez que a exposição ao *media* nem sempre se encontra condicionada pelos conteúdos transmitidos. Quando se consideram os programas que apresentam uma maior regularidade de valores dos indicadores em análise, e que nem sempre são transmitidos fora dos períodos de maiores audiências, os resultados das análises de previsão podem ser mais satisfatórios. Para os programas mais estáveis poderá não ser necessário desenvolver modelos de previsão, quer dizer, podemos simplesmente considerar o cálculo de médias de valores históricos considerando apenas esse programa ou em função da subamostra de programas que apresente maiores semelhanças face a esse conteúdo e quanto às características da contra programação e da programação adjacente.

Todavia, inclusive para o período matinal e da tarde, onde os indicadores apresentam uma evolução mais regular em dias laborais ou ao fim de semana, também foram realizadas análises de duplicação de audiências, para se verificar qual é o grau de estabilidade nos padrões de exposição televisiva que lhe são subjacentes.

Verificamos que a regularidade dos indicadores não se associa necessariamente a uma elevada duplicação das audiências. Neste capítulo pretende-se responder a uma das questões apresentada no *capítulo I* e que se transcreve no seguinte parágrafo:

Questão 3: “Qual a viabilidade de construção de planos de media, antes de ser aplicado o processo de optimização, com base na Previsão de valores de “ratings” programáticos e comerciais, recorrendo a metodologias Multivariadas, Simbólicas e de Data Mining proposicional?”

Foram privilegiadas as análises multivariadas, que, também, podem apresentar potencial predictivo, e em especial as análises não lineares. Como já foi referido anteriormente, e internacionalmente demonstrado, quando se recorre aos inquéritos de intenção de ver os resultados não são satisfatórios e as análises incidem nos períodos de maiores audiências uma vez que estes processos são onerosos e apenas justificáveis nos períodos que maximizam as receitas publicitárias e que apresentam uma maior variabilidade.

A previsão qualitativa, também, se deveria utilizar nas análises de previsão dos suportes que correspondem a novos programas e quando se transmitem “eventos” com uma influência muito variável ao nível da evolução das taxas médias de audiências para o total da TV e da sua capacidade em conseguir captar os telespectadores de canais concorrentes. Nestes contextos em *Nikolopoulos et al. (2006)* com dados de 46 eventos desportivos, transmitidos de 1996 a 2000, os autores, também, recorrem a abordagens quantitativas, nomeadamente a determinadas metodologias de *Data Mining*, como, por exemplo, aos algoritmos KNN e a Redes Neurais, a metodologias de Regressão linear simples e múltipla e a Juízos de Valor. Os autores concluem que os melhores resultados resultam das análises de Regressão linear Simples e dos algoritmos KNN que apenas realizam uma média de um conjunto de observações mais semelhantes à observação a prever segundo os critérios considerados nas análises. Salientam em relação aos métodos qualitativos que:

Judges also have limited information processing capacity, which often causes them to over simplify problems, to produce biases judgments and to behave inconsistently.... in application of judgment to time series data, human forecasters often display a characteristic similar to overfitting in that they see imaginary systematic patterns in noise and attempt to extrapolate these patterns when producing their forecasts. (p.365)

Da amostra de treino apenas constam 34 observações e 12 na amostra de teste, sendo especialmente difícil encontrar observações muito semelhantes se forem considerados vários critérios. Relativamente à robustez dos sistemas, ou seja, ao seu comportamento perante a variabilidade que se verifica no panorama audiovisual, por vezes, é referido que os sistemas que funcionam por análise de situações passadas e que promovem a identificação de padrões existentes, ainda que mais ambiciosos tendem a ser na realidade muito pouco realistas.

As análises recorrendo a séries temporais devem ser consideradas nos períodos ou contextos de maior estabilidade e regularidade dos indicadores, que também foram identificados nas análises realizadas. Porém, apresentam um menor “interesse” para quem desenvolve a actividade de planeamento de *media*. Na opinião de *Danaher et al. (2011, p.22)*, inclusive quando se recorre à análise das séries temporais de *ratings* devem ser considerados os dados multivariados: “It is always important to include exogenous covariates, such as holidays, season, day of week, time of day and yearly trends”.

Em qualquer área, quando se dispõe de suficiente informação estatística e de uma previsão relativamente precisa dos factores externos ou das variáveis explicativas, devemos recorrer aos modelos causais, sendo esta a perspectiva privilegiada nesta tese, ainda que não a utilização das séries temporais multivariadas, uma vez que se pretendeu a análise de um período característico a nível médio do consumo televisivo e a avaliação da capacidade explicativa de determinados comportamentos e variáveis. Também, se privilegiou nas análises, o horizonte temporal que corresponde a um maior dinamismo neste sector, avaliando-se assim qual a sua repercussão na evolução dos dados e na viabilidade de serem realizadas análises de previsão.

Verificamos que apesar das consideráveis alterações observadas nos *shares* de audiências dos diversos canais generalistas, nomeadamente a consolidação da posição da TVI e uma perda de *share* da SIC, uma vez que estamos perante um oligopólio e o painel na sua maioria não apresenta acesso à TV Cabo, alguns dos indicadores não são muito mais variáveis do que quando se considera períodos de maior estabilidade, como, por exemplo, as taxas médias de audiência para o total da TV.

Uma vez que no decorrer dos anos, por vezes, se verifica uma reestruturação parcial do painel com a substituição de determinados indivíduos ainda que, mantendo a sua composição sociodemográfica, as alterações observadas, também, podem decorrer destas alterações.

Nas análises privilegiaram-se as análises não lineares e as análises simbólicas, recorrendo a diferentes metodologias de *Data Mining* proposicional e a variáveis explicativas que constam das bases de dados de audimetria, assim como a alguns escassos algoritmos que visam a análise das séries temporais de valores de *ratings* e de taxas médias de audiências para o total da TV, apenas para avaliar a viabilidade de recorrer a estas abordagens e qual a dimensão dos erros de previsão. Estes algoritmos cumprem os requisitos que os sistemas de previsão apresentam.

A instabilidade verificada, por vezes, ao nível dos programas e das tipologias transmitidas, estratégias de contra-programação e as alterações horárias poderia ou não dificultar a realização das análises de previsão. Em inúmeros contextos, para o período em análise, as alterações a este nível não influenciam, fortemente, os resultados, desde que considerando a evolução dos indicadores de audimetria. Contudo, a evolução agregada não corresponde ao somatório das decisões individuais, como facilmente se constatou com a análise dos valores de duplicação em contextos de maior estabilidade de *ratings* em que se pretendeu associar a evolução dos *ratings* com a evolução dos índices de duplicação das audiências.

Verificámos que poderá ser viável a previsão de *ratings* em determinados contextos, uma vez que os coeficientes de determinação podem atingir os 80% considerando, quer as análises de Regressão linear, quer as análises de Regressão não linear nos períodos de maiores audiências, superando os valores que constam da bibliografia internacional da especialidade e que incidem em bases de dados mais recentes. Contudo, os valores de coeficientes de determinação que se aproximam dos 80% podem ser considerados satisfatórios, tendo em atenção os resultados que se encontram publicados, mas certamente que o sector televisivo precisa alcançar menores margens de erro.

Por exemplo, em *Weber (2002)* com valores de erro ligeiramente superiores, o autor conclui que esses valores são satisfatórios, atendendo às dificuldades que estas análises têm colocado ao sector durante décadas. Porém, previamente será necessário conseguir prever os valores das taxas médias de audiências para o total da TV.

As análises realizadas a *Tvr%* apresentam margens de erro que podem ser consideráveis, mas outras metodologias podem ser utilizadas e devemos considerar dados para menores unidades temporais, apresentando-se no seguinte capítulo algumas sugetões.

Para promover a previsão das taxas médias de audiências para o total da TV recorreu-se a análises de previsão e de segmentação. Relativamente às análises de previsão recorrendo a metodologias “Nearest Neighbour”, também, denominadas de “Instance Based Learners” (IBL), a Regressão Kernel, a Regressão Polinomial e a Regressão Local linear e Local Polinomial conseguiu-se no máximo uma redução no erro quadrático médio em 75,7% para os dados semanais. Atendendo aos dados de fim-de-semana, as percentagens são inferiores e no máximo com o valor de 58,7%. Os melhores modelos correspondem a Regressões ultralocais lineares e quadráticas. Atendendo à influência da proximidade temporal nestes valores, recorrendo-se aos IBL avalia-se o grau de semelhança existente entre os valores que são temporalmente próximos.

Sem dúvida, que existem interdependências, constrangimentos que correlacionam o valor em determinado instante com os valores para os instantes posteriores, mas, dificilmente, as estações televisivas conseguem manter os seus índices de audiências ao longo de um maior horizonte temporal. A utilização de Redes Neurais MLP e das Médias Móveis não acrescentou grandes mais-valias às análises, apresentando percentagens de variância explicada com a mesma ordem de grandeza.

Através das análises de segmentação conseguimos, em função da variabilidade observada nos dados, a subdivisão da emissão diária em períodos comparáveis. Alguns algoritmos simbólicos apresentam percentagens de inércia explicada que ultrapassa os 90%. A emissão diária subdivide-se em períodos com evoluções lineares e polinomiais.

Quanto mais acentuados forem os efeitos de herança observados entre os programas, mais semelhantes podem ser os valores da variável em estudo para observações temporalmente mais próximas. No entanto, os efeitos de herança por si só são insuficientes, uma vez que esta variável também depende das decisões de ligar e desligar o televisor e dos fluxos de audiências entre os canais. Os maiores valores de efeitos de herança no mínimo impedem uma maior redução dos valores desta variável para períodos mais próximos. Face aos resultados das análises realizadas com o intuito de prever a variável em estudo, novas abordagens metodológicas devem ser implementadas e diferenciadas em função dos contextos e estabilidade dos dados de audimetria.

Os melhores resultados decorreram das análises de segmentação dos dados e podem produzir melhores resultados se considerarmos séries temporais ao segundo, levando a cabo as necessárias análises de pré processamento dos dados. As análises de segmentação tiveram diferentes objectivos, entre os quais, a viabilização da aplicação de determinadas metodologias com esse intuito.

Alguns dos objectivos foram os seguintes:

- A Transformação simbólica da série temporal original. Atendendo aos resultados alcançados, nem sempre os clusters horários contêm faixas horárias sequenciais. Os períodos temporalmente distantes, por vezes, apresentam evoluções comparáveis. Através das análises simbólicas podemos associar símbolos nominais aos períodos horários, viabilizando assim a posterior aplicação e derivação de Regras de Associação ou de Regressão.
- Identificação de faixas horárias comparáveis quanto à evolução desta variável, uma vez que os métodos de previsão mais adequados variam no decorrer da emissão diária e são função da amplitude de valores observados.

Para além de auxiliar na identificação das subamostras temporais dos dados, também permite a discretização da variável em estudo, viabilizando a aplicação de metodologias que não seriam viáveis com dados numéricos.

A discretização traduz-se na conversão de um valor numérico em um valor nominal ordenado, ou seja, num ordinal. Existem infinitas possibilidades quanto aos pontos extremos dos intervalos considerados e amplitude dos mesmos, consequentemente adoptaram-se várias metodologias para se comparar os resultados. Posteriormente, para avaliar a qualidade da discretização podemos avaliar a sua repercussão nos valores de *Ratings* e *Shares*, por exemplo, por aplicação das análises ANOVA. Perante uma amostra de dados associada a uma grelha de programação normalmente transmitida, nas análises de previsão das taxas médias de audiências para o total da TV promoveu-se a aplicação de metodologias de Regressão não paramétricas. Privilegiou-se a modelização não paramétrica e em particular a que faz uso de métodos de núcleo (Kernel) para poder construir modelos mais flexíveis uma vez que se ajusta aos dados localmente.

Os métodos paramétricos apresentam como desvantagem a rigidez da sua estruturação não se adaptando a inúmeras amostras de dados. Os resultados são mais satisfatórios para o período semanal, no entanto, não transcendem os 75% de variância explicada.

No âmbito não paramétrico existe um grande número de metodologias: ajustes locais de modelos paramétricos já utilizados, os métodos baseados em séries ortogonais de funções, Splines e técnicas de aprendizagem supervisionada que podem ser implementadas em análises posteriores.

Em análises a realizar, uma vez que a via do *clustering* de séries temporais apresentou os melhores resultados, será necessário recorrer a outras metodologias do âmbito do clustering e classificação de ST adoptando outras medidas de avaliação de similitude e de representação das séries. No âmbito do Indexing podemos recorrer aos Vp-Trees, Hb-Trees ou aos Grid-files e promover a aplicação de inúmeras medidas de dissimilaridade existentes entre séries univariadas e multivariadas.

Para os segmentos que condicionam em maior grau a evolução deste indicador e que apresentam uma maior variabilidade podemos recorrer a dados dos ficheiros base de audimetria e utilizar as Regras Sequenciais para prever, não só, as sequências de selecções programáticas mais prováveis, mas, também, as decisões de ligar e desligar o televisor.

Os principais factores que condicionam a variabilidade de *ratings* são a notoriedade do canal, os dias (semana versus fim de semana), faixas horárias, alterações horárias e a programação desportiva. No período matinal, também, são viáveis as análises de previsão de *ratings* não sendo, frequentemente, necessário recorrer a variáveis explicativas, uma vez que a análise das séries temporais costuma produzir bons resultados.

Desde que previstos os valores dos *ratings* programáticos facilmente podem ser calculadas as audiências comerciais, atendendo à elevada correlação linear que existe entre ambos os indicadores. Verifica-se uma maior dificuldade na identificação dos factores responsáveis pelas quedas das audiências comerciais e na sua previsão, quantificadas através do indicador *Spot_Ratio%*.

Em contextos de maior estabilidade de valores de *ratings*, as previsões podem ser realizadas com base na análise de séries temporais desde que os canais não realizem significativas alterações na grelha de programação e se excluam determinados eventos desportivos. Verificou-se que alternativamente a via da segmentação e o clustering de séries temporais também viabiliza resultados satisfatórios, face a uma grelha que não apresente programas de maior “notoriedade”. Por vezes, estas análises podem ser realizadas ignorando a programação dos canais concorrentes.

O grau de variabilidade que se observa nos *ratings* permite colocar dois cenários, consoante é necessário conhecer a programação dos diversos canais em determinada faixa horária ou não para realizar a previsão. Quando não se conhece com precisão a programação a solução mais fiável será extrapolar o desempenho passado de determinado canal.

Porém, ainda que existam alterações, por vezes, os programas alteram a programação, mas substituem uns programas por outros de peso equivalente para manterem os seus níveis de audiências. Logo, as extrapolações, nestes contextos, também podem ser consideradas. A nível nacional, uma vez que se recorre, frequentemente, à transmissão de séries, telenovelas, “talk-shows” e “reality shows” a previsibilidade dos conteúdos é superior. Em contextos de maior estabilidade dos dados, a extrapolação a partir do passado e outras abordagens de análise de séries temporais permitem alcançar resultados satisfatórios.

Todavia, devem-se privilegiar as abordagens que não estão associadas à previsão das séries temporais, mas, sim, de outros âmbitos, dado que a sua utilização em horizontes temporais de maior dimensão é problemático uma vez que a essência da série não está na sua tendência. As audiências variam no decorrer do tempo, por exemplo, como consequência da sua alteração horária quando se referem a valores programáticos ou do êxito de um programa transmitido na mesma faixa horária na programação concorrente ou da variabilidade associada à exposição ao *media* e da permanência da audiência no decorrer dos programas.

Logo, não se deve procurar a tendência como o fazem alguns métodos quantitativos de previsão, como, por exemplo, os métodos de decomposição ou de *Box Jenkins*. Quando consideramos um maior horizonte temporal devem ser proporcionadas variáveis explicativas, exógenas e que viabilizem também uma interpretação dos dados face aos modelos meramente econométricos univariados. A utilização de metodologias simbólicas constitui uma opção, assim como o *Data Mining* temporal.

Certamente, que o desempenho de um programa encontra-se influenciado pela estacionalidade. Um valor mais reduzido dos *ratings* em determinado mês poderá resultar de que nesse mês se vê menos televisão e não de uma diminuição da eventual “atractividade” que determinado programa exerce sobre o painel ou de significativas alterações nas restantes influências estruturais. As análises univariadas de séries temporais prolongam uma tendência e têm em conta as variações estacionais. A prolongação da tendência realiza-se segundo uma linha recta, ou não, e por vezes, tendo em conta em maior grau o passado mais longínquo.

Alguns algoritmos atribuem maiores factores de ponderação às observações mais recentes, mas para se aplicar esta tipologia de abordagens será necessário considerar um horizonte temporal com alguma amplitude, o que não constitui uma limitação uma vez que nas bases de audimetria consta a informação relativa aos últimos cinco anos. Se apenas tivéssemos em conta o passado mais recente não existiria uma série temporal. Se os dados não estiverem disponíveis podemos recorrer a metodologias, por exemplo, de *clustering* de séries temporais, entre outras do *Data Mining* Temporal.

O passado mais remoto, frequentemente, não é considerado como realmente relevante para a análise das audiências, uma vez que se verificam alterações no cenário da comunicação social e ao nível da TV que podem influenciar fortemente os resultados. Para os vários anos considerados nas análises, verificamos que apesar das alterações verificadas no sector, por vezes, os conteúdos transmitidos não influenciam fortemente a evolução das taxas médias de audiências para o total da TV.

Verifica-se uma maior alteração nos valores dos shares de audiências e que, por vezes, também, se reflecte em menor grau na evolução dos *ratings* que corresponde ao indicador que realmente condiciona a construção dos planos. As previsões de *ratings* também podem ser realizadas com base na prévia análise das taxas médias de audiências para o total da TV e dos valores de *shares*.

As análises meramente exploratórias de previsão de *Tvr%* apresentam margens de erro que podem ser muito consideráveis, independentemente das abordagens que foram adoptadas. Para tentar prever os valores de shares podemos recorrer à extrapolação que não tem em conta a programação dos canais concorrentes ou recorrer a análises multivariadas que o considerem. Contudo, este indicador apresenta para o período em análise uma forte variabilidade, sendo preferível considerar directamente os valores dos *ratings* ou em alternativa apenas por associação com as taxas médias para o total da televisão.

Neste trabalho não foram desenvolvidos modelos explicativos da evolução dos *shares* de audiências. A publicação mais recente identificada nesta área, também, refere não ser necessário recorrer aos valores de shares quando se pretende desenvolver modelos de *ratings*. “.... our finding regarding the superiority of forecasting *ratings* directly is contrary to the established wisdom, which recommends the development of separate models for the total audience and channel share” *Danaher et al. (2011, p.22)*.

As previsões devem ser realizadas para cada canal, dia e faixas horárias de 15 ou 30 minutos. Nas análises realizadas considerou-se cada um dos canais e diversos períodos horários. Os dados de *Tvr%* e de *Rat%* que se referem a unidades temporais foram calculados para períodos de 30 minutos. As análises devem ser realizadas para unidades de tempo que se aproximem das durações dos conteúdos, fazendo corresponder os programas a conjuntos de unidades temporais, o que poderá facilitar a extrapolação dos resultados para o contexto programático. Segundo *Danaher et al. (2011, p.22)*: “We found developing separate models for each channel results in much better forecasts than having a combined model with channel dummy variables”.

Para as previsões de longo prazo devemos considerar maiores intervalos de tempo, uma vez que será mais difícil realizar a previsão, mas considerando espaços horários que se relacionem como o modo de vida dos telespectadores e com a estruturação das grelhas. As análises de segmentação das séries temporais são relevantes neste contexto e constam das análises realizadas. Devem-se prever as taxas de audiências em função dos dias laborais ou de fim-de-semana e utilizar os índices de estacionariedade para diferenciar a audiência dos diferentes períodos do ano.

Na estacionalidade os meses estivais apresentam os menores valores e os meses de inverno têm maiores valores. Portanto, devem existir factores correctores para cada mês, por períodos do dia e por públicos objectivos. Os vários segmentos podem ser agregados em função da evolução dos dados, tendo-se avaliado este aspecto nas análises, considerando variáveis da tipologia intervalo e as metodologias simbólicas.

Nas análises realizadas, também, se avaliou a repercussão das variáveis sociodemográficas na evolução dos *ratings* e das taxas médias de audiências para o total da TV, existindo determinados segmentos que se diferenciam significativamente da evolução mais característica para o universo televisivo e que, também, devem ser identificados para cada canal e segmento horário quando se pretende realizar previsões de *ratings* para o universo televisivo por ponderação das previsões dos *ratings* realizadas para cada um dos segmentos (Meyer & Hyndman, 2005). Nas análises das tendências diferenciam-se as tendências de longo prazo das audiências médias para o total da TV por faixas horárias, que se assume até certo ponto maioritariamente independente dos conteúdos transmitidos, das tendências a curto prazo dos *ratings* dos canais por faixas horárias. Os modelos de previsão devem excluir a informação histórica associada a grandes acontecimentos. A existência de alguma estabilidade ao nível das taxas médias para o total da TV corresponde a um ponto de grande relevância e que constitui um ponto de apoio em circunstâncias de comportamentos voláteis e dinâmicos, contudo determinados algoritmos não apresentam resultados muito satisfatórios para determinados contextos de análise.

A regularidade que por vezes se observa ao nível das taxas médias para o total da TV não implica necessariamente a existência de regularidade na exposição televisiva individual e agregada para a totalidade do painel.

Nos períodos com audiências mais elevadas e variáveis, também, podem ser realizadas análises recorrendo a Árvores e a Regras de Regressão de *Data Mining* que correspondem a sistemas que interpretam os dados e que produzem um conjunto de conhecimento fundamentado e extraído dos dados. As Regras são inteligíveis para os utilizadores e identificam informação que se encontra oculta nas bases de dados sob a forma de padrões, podendo ser realizadas análises de previsão ou de classificação. Estes algoritmos produzem Regras que são o resultado de um processo de indução. Em comparação com as Redes Neurais apresentam a vantagem de não ter um funcionamento opaco, com saídas explícitas e facilmente interpretáveis, o que permite aos utilizadores aceitar ou rejeitar determinados resultados.

Entender melhor o comportamento da audiência e dos indicadores tem certamente muito valor para os programadores, ajudando na identificação de oportunidades num meio que se encontra em constante evolução. Qualquer resultado que decorra de um método quantitativo ou qualitativo de previsão corresponde apenas a um dado que o utilizador final aceitará ou alterará em função da sua experiência e intuição. Os dados de audimetria a prever também podem ser discretizados e estabelecer-se uma regra de classificação para cada um dos intervalos de valores.

Posteriormente, será necessário verificar se uma nova observação corresponde ou não a esse intervalo. Estes algoritmos englobam as seguintes fases: Selecção dos aspectos que devem ser considerados para produzir as previsões mais exactas, como, por exemplo, a informação do programa, mas também dos canais concorrentes e dos programas essencialmente que os antecedem e dos quais poderá “herdar audiência”; processamento dos dados; aplicação dos sistemas de aprendizagem aos dados de treino; análise de resultados; validação recorrendo a amostra de teste.

As variáveis que foram identificadas como possíveis variáveis de previsão de valores de *ratings* correspondem às variáveis temporais, por exemplo, o período horário, horas de transmissão dos suportes e as taxas médias de audiências para o total da TV, à tipologia do programa, ao canal, a fidelidade programática e os valores dos principais indicadores de audimetria para os programas adjacentes, assim como alguns reduzidos factores de planeamento associadas a durações e a saturações publicitárias.

Seleccionando as variáveis mais relevantes, calculam-se as previsões com base em Regras que definem que para determinado conjunto de programas, com determinadas características produzem-se determinadas audiências em determinados canais. Verificamos que a nível nacional os factores temporais são determinantes. Quando se tem determinadas Regras para se realizar a previsão será necessário conhecer a programação que vai ser transmitida no período em análise, o que a nível nacional é viável em grande parte do ano, uma vez que são transmitidas inúmeras séries e telenovelas.

Quando se alteram as grelhas, por vezes conhecendo o modo de proceder de quem define os conteúdos a transmitir não será possível dizer qual o programa em concreto que vai ser transmitido, mas, sim, a sua tipologia. Por vezes, a estabilidade dos conteúdos decorre da estratégia do canal e, conseqüentemente, acabamos por não encontrar uma grande variabilidade a este nível. Com base nas análises multivariadas realizadas, a utilização de metodologias de *Data Mining*, como, por exemplo, as Redes Neurais ou as Árvores e as Regras de Decisão permitiram atingir resultados que podem ser considerados como razoáveis para uma grelha de programação já estabelecida, atendendo aos valores que constam da bibliografia internacional, mas provavelmente insuficientemente precisos para as necessidades deste sector.

A percentagem de variância explicada pode superar os 80% para a maior parte dos canais e períodos considerados, mas pode implicar a necessidade de serem previstos algumas das variáveis, nomeadamente os valores dos *ratings* para os suportes adjacentes o que poderá reduzir a capacidade predictiva destes modelos.

Se apenas forem consideradas variáveis que são conhecidas no momento de realizar uma previsão os coeficientes de determinação são menores. Segundo as análises não lineares e por aplicação do CHAID à totalidade da amostra a percentagem de variância explicada é de 87% para amostra de teste e as variáveis utilizadas no desenvolvimento da Árvore correspondem a *Tvr_Prog%*, tipologias programáticas, canal e os *ratings* de programas adjacentes. O algoritmo CART, que do ponto de vista da previsão é mais adequado, apresenta 82% de variância explicada e incide nas mesmas variáveis.

Se restringirmos as análises apenas aos canais privados para o algoritmo CART a percentagem de variância explicada é de 85% e para o canal estatal RTP1 de 80% incidindo essencialmente em variáveis temporais, tipologias, canal e a fidelidade programática. Consequentemente, a capacidade predictiva destes modelos poderá ser insuficiente atendendo à precisão necessária quando se constroem os planos de *media*. Recorrendo às Redes Neurais MLP e RBF verificamos que discretizando os valores de *ratings* em intervalos de 4%, 82,13% das observações ficam correctamente classificadas para as Redes Neurais MLP e 80% para as Redes Neurais RBF.

Estes valores são até certo ponto equiparáveis aos que resultaram das análises realizadas com as Árvores de Regressão. As metodologias simbólicas, também, apresentam valores dentro desta ordem de grandeza, por exemplo, a SIC e a RTP1 aproximam-se dos 80% e para a TVI poderá ser ligeiramente superior. Porém, melhores resultados podem ser alcançados se considerarmos os valores referentes a períodos horários em vez dos valores programáticos. A nível nacional a estrutura horária é determinante em análises de previsão, mas sem dúvida que são necessárias abordagens alternativas de construção dos planos de *media* que permitam ultrapassar a necessidade de prever os valores de *ratings*.

Os principais indicadores em análise devem ser os valores dos *ratings* e as taxas médias para o total da TV. As restantes variáveis a considerar nos modelos “causais” correspondem às faixas horárias e a outras variáveis de cariz temporal, as tipologias programáticas e a caracterização da programação envolvente, sendo mais relevante para

grande parte do período em análise a programação adjacente e pontualmente a concorrente com uma repercussão muito variável no desempenho de cada um dos canais e na evolução dos indicadores de audimetria. Em *Danaher et al. (2011)*, também apenas foram utilizadas as variáveis de audimetria. Consideraram as variáveis temporais, tipologias, duração do programa e se consiste em uma retransmissão e se a tipologia coincide com o histórico para essas faixas horárias e a tipologia de “Lead_In”.

Para melhorar os resultados das análises de previsão podemos aumentar a dimensão dos dados utilizados nas bases de dados de treino, melhorar a selecção e a especificação dos dados, assim como considerar o maior número possível de variáveis de caracterização da grelha de programação e do painel e que não constam das bases de audimetria.

Nas análises realizadas foram considerados todos os indicadores, as variáveis de planeamento e de caracterização da grelha que constam das bases de audimetria associadas ao *software* Telereport e Publitex. As previsões podem ser realizadas para faixas horárias, programas ou tipologias e por públicos objectivo ou para o universo televisivo.

Nem todos os segmentos sociodemográficos apresentam igual repercussão na evolução dos índices de audiências, colocando maiores obstáculos à previsão, os que apresentam uma maior irregularidade na exposição. Consequentemente, realizaram-se análises que permitiram identificar quais são os segmentos que se aproximam em maior grau do universo e entre si, constituindo macro-segmentos de análise para os diversos canais e faixas horárias. Os *ratings* também podem ser previstos por ponderação dos valores associados a cada um dos segmentos.

Face à transmissão de programas que geram observações atípicas, a previsão poderá colocar maiores obstáculos, uma vez que se encontra, fortemente, dependente da exposição dos telespectadores que não apresentam regularidade na sua exposição ao *media* e a influência do conteúdo poderá ser superior.

Facilmente, identificamos uma grande diversidade de cenários que resultam da transmissão de eventos desportivos e que por vezes não influenciam a evolução dos *ratings* apesar das fortes alterações que se verificam nos shares de audiências e que são consequência das alterações que se verificam na exposição global ao *media*.

Do ponto de vista matemático o desenvolvimento de modelos explicativos e de previsão das audiências comerciais corresponde a uma questão que assume uma grande complexidade devido aos seguintes factores:

- A audiência tem que ser prevista para níveis muito detalhados de análise, sendo necessário considerar os valores de cada inserção e não o espaço comercial na sua totalidade. Por vezes é conveniente a quantificação das audiências ao segundo;
- As audiências comerciais estão fortemente dependentes das audiências programáticas desse canal e, por vezes, mas em menor grau dos canais concorrentes, os quais nem sempre divulgam a grelha de programação com a necessária antecedência e de forma fiável. Porém, a nível nacional, incidindo a programação em séries e em telenovelas nos períodos de maiores audiências permite-nos, até certo ponto, ultrapassar esta limitação;
- As previsões devem ser realizadas com rapidez, uma vez que condicionam o valor das tarifas e a venda de espaço comercial. Logo, devem ser avaliadas as alternativas viabilizadas pelas bases de audimetria e que constam da presente tese;
- As previsões têm que ser realizadas para diferentes públicos objectivos que podem apresentar divergências quanto à forma de interacção com o meio televisivo. Porém, a evolução global não corresponde a uma mera agregação dos comportamentos individuais e por segmentos. Neste trabalho, também, se avalia a repercussão destas variáveis na evolução dos *ratings*;
- As previsões têm que ser realizadas para diferentes faixas horárias, canais, tipologias programáticas, entre outras variáveis. Existe uma forte diversidade de contextos explicativos e de previsão, cada um dos quais exigindo diferentes abordagens analíticas e/ou fontes de dados. Procurou-se avaliar em que medida as variáveis estruturais condicionam a evolução deste indicador;

➤ A previsão apresenta grandes obstáculos no âmbito das Ciências Sociais, uma vez que múltiplas variáveis que não são quantificáveis e imprevisíveis influenciam a exposição e os comportamentos futuros dos telespectadores. Consequentemente, avalia-se a capacidade explicativa e, eventualmente, de previsão dos modelos lineares e não lineares de audimetria e recorre-se aos dados agregados, uma vez que são os mais relevantes para o sector e os que podem permitir a identificação de regularidades nos dados viabilizando a sua previsão. Ao contrário, do que se verifica nas Ciências Naturais, onde, por vezes, é viável a previsão, recorrendo a um modelo matemático e considerando um reduzido número de variáveis quantificáveis e conhecidas ou estimáveis, quando estão em análise comportamentos surgem maiores dificuldades, sendo, frequentemente, inviáveis.

No contexto das inserções comerciais e considerando apenas os valores dos *ratings*, as análises de previsão podem ser, eventualmente, viáveis ainda que com determinadas margens de erro que não devem ser consideradas como excessivamente grandes, atendendo aos resultados que constam da bibliografia da especialidade. Previamente, têm que ser calculadas as audiências programáticas que lhe estão associadas e que, por vezes, podem colocar maiores obstáculos quanto à viabilidade da sua previsão.

Os modelos de Regressão linear múltipla apresentam percentagens de variância explicada que podem atingir os 82% considerando a amostra global dos canais estatais e de 87% para os canais privados generalistas, existindo uma elevada associação linear entre os valores dos *ratings* programáticos e os *ratings* comerciais. Dado que os blocos comerciais não apresentam uma dinâmica própria e independentemente dos conteúdos programáticos que lhe estão associados, a previsão das audiências programáticas pode viabilizar a previsão das audiências das inserções publicitárias que na realidade apresentam uma forte evolução no decorrer dos espaços comerciais.

Para o canal SIC a associação linear entre os *ratings* programáticos e os *ratings* comerciais para a totalidade da amostra apresenta um coeficiente de determinação no valor de 0,79 e em alternativa se incluirmos mais duas variáveis de planeamento que são

conhecidas no momento de realizar a previsão, nomeadamente a duração de um programa e o número de promoções que integram um espaço comercial, o valor do coeficiente de determinação é de 0,80. Se associarmos os *ratings* programáticos com a hora de transmissão da inserção o valor do coeficiente de determinação é de 0,82. Consequentemente, correspondem a valores que podemos considerar razoáveis atendendo aos resultados que constam da bibliografia internacional.

Para o canal TVI considerando apenas os *ratings* programáticos o coeficiente de determinação é de 0,85 e se adicionarmos as taxas médias de audiências para o total da TV e as durações acumuladas das inserções o valor do R^2 ascende a 0,89. Se considerarmos apenas a variável *Tvr%* este valor seria apenas de 0,57. Portanto, os modelos lineares podem apresentar alguma capacidade de previsão desde que previstos os principais indicadores de audimetria programáticos, aos quais se associam determinadas variáveis de planeamento. Para o canal RTP1, e considerando as mesmas variáveis referidas para o canal TVI o coeficiente de determinação é de 0,82.

Considerando o indicador *Spot_Ratio%* que quantifica a queda das audiências comerciais, mas a nível relativo face aos valores dos *ratings* programáticos que lhe estão associados os resultados são consideravelmente menos satisfatórios. A capacidade predictiva dos modelos lineares considerando as amostras globais é irrisória para os canais estatais e privados.

Considerando apenas a amostra global do canal SIC, o coeficiente de determinação máximo alcançado com as variáveis de audimetria é de 0,22 e representa um modelo que inclui os *ratings* programáticos e algumas variáveis conhecidas quando se realiza a previsão, nomeadamente a hora de transmissão da inserção, a duração acumulada desde o início do espaço comercial e a transmissão da inserção, a duração do programa e a hora de conclusão do programa.

Consequentemente, não é viável a previsão deste indicador recorrendo aos modelos de RLM. Se restringirmos as análises às faixas horárias da noite o coeficiente de determinação não transcende os 0,43 para o canal SIC. Para o canal TVI, também, se verifica uma ausência da linearidade das relações entre as variáveis de audimetria. Para a totalidade da amostra deste último canal o melhor coeficiente de determinação corresponde a 0,18 e a um modelo que inclui o rating do programa, a taxa média de audiência para o total da TV, a duração acumulada desde o início do espaço comercial e a transmissão da inserção e a saturação publicitária do programa. Consequentemente, também, não apresenta uma significativa capacidade predictiva, verificando-se a mesma conclusão para o canal RTP1 apresentando o maior coeficiente de determinação o valor de 0,25.

Verificamos que as variáveis que condicionam a evolução dos *ratings*, absolutos ou relativos, diferenciam-se para os vários canais generalistas, sendo as mais relevantes a posição da inserção publicitária e a localização do espaço comercial, assim como a faixa horária considerada, os principais indicadores de audimetria e as tipologias programáticas envolvidas. Observa-se a existência de significativas interações entre as variáveis mencionadas e que dependem do canal considerado.

Segundo as análises simbólicas de Regressão linear e considerando a variável *Rat_Spot%* foi possível identificar um modelo com um coeficiente de determinação no valor de 0,87 e que inclui apenas as variáveis que podem ser conhecidas *a priori* ou calculadas no momento de realizar a previsão, por exemplo, recorrendo a *Tvr_Prog%* que poderá apresentar uma menor variabilidade nos seus valores para determinados horizontes temporais.

Assim como se verificou para os modelos de Regressão linear não simbólica, o valor da percentagem de variância explicada pode ser considerado como sendo razoável e apresenta alguma capacidade de previsão, atendendo às variáveis que foram consideradas nas análises. Se excluirmos do modelo anterior as variáveis associadas às taxas médias de audiências para o total da TV obtemos um R^2 de 0,80.

Porém, uma vez mais para a variável *Spot_Ratio%* se considerarmos apenas as variáveis conhecidas ou que podem ser calculadas previamente conseguimos um R^2 no valor de 0,39, ou seja, muito menor ao do modelo associado a *Rat_Spot%*, mas superior aos dos modelos de Regressão linear tradicionais.

De acordo com as análises de Regressão não linear realizadas aos valores dos *ratings* comerciais e por aplicação do algoritmo CART, a percentagem de variância explicada para a amostra de teste é de 94%, considerando os dados do canal RTP1. As variáveis que se associam com este indicador correspondem aos *ratings* programáticos, aos períodos horários e à duração desde o início do programa até à transmissão da inserção. Se excluirmos das análises os *ratings* programáticos verificamos que os níveis de audiências comerciais são função de *Tvr_Prog%* e das faixas horárias, apresentando este modelo 85% de variância explicada.

Consequentemente, as análises não lineares apresentam alguma capacidade de previsão e superior às dos modelos lineares desde que, também, seja viável prever os valores dos principais indicadores de audimetria programáticos. Assim como se verificou para os modelos lineares a variância explicada quando se considera a variável *Spot_Ratio%* fica muito aquém da calculada para os valores dos *ratings*. Para a variável *Rat_Spot%* do canal SIC a variância explicada é no máximo de 90% e uma vez mais destacam-se os factores temporais e os principais indicadores de audimetria. Conduzindo idênticas análises com a variável *Spot_Ratio%* uma vez mais a percentagem de variância explicada não atinge os 50% considerando os dados para ambos os canais privados.

Com base nas análises lineares realizadas verificamos que quando consideramos os *ratings* comerciais e em especial as análises simbólicas constituem uma alternativa às análises de Regressão não linear baseadas em Árvores de Regressão, apresentando ambas as metodologias razoáveis percentagens de variância explicada para os diversos horizontes temporais em análise. Existe uma forte associação linear e, também, não linear entre os *ratings* programáticos e os *ratings* comerciais.

Contudo, resultados muito diferentes foram alcançados para o indicador *Spot_Ratio%*, para o qual as análises de Regressão linear e não linear não permitiram alcançar resultados satisfatórios que possam viabilizar a previsão deste indicador e a identificação das variáveis que condicionam a sua evolução.

Os índices comerciais são função dos índices dos programas em que se encontram inseridos, mas ficou demonstrado que não correspondem simplesmente ao preenchimento da audiência entre dois programas. A perda de telespectadores no decorrer dos espaços comerciais pode e deve ser calculada para alguns meses. Devem-se calcular estes factores correctores para cada canal, fracção do dia que resulte das análises de segmentação e por público objectivo.

Nas análises realizadas, também, se avaliou a repercussão das variáveis sociodemográficas na evolução dos *ratings* comerciais. Os espaços comerciais têm que ser diferenciados em função de serem transmitidos no decorrer de um programa ou entre programas e foram identificadas algumas variáveis de planeamento que condicionam em maior grau a sua evolução. A identificação de quais são os factores que se relacionam com o “factor bloco” é de grande relevância para o sector e, consequentemente, também se realizaram algumas análises explicativas. Contudo, não foi possível, recorrendo apenas às variáveis de audimetria, identificar “causas” ou associações que permitam explicar a sua evolução. Em suma, uma vez previstas as audiências programáticas podemos calcular as audiências comerciais, atendendo à significativa relação linear e não linear observada entre ambas as variáveis.

Do ponto de vista comercial não se deve realizar a previsão das audiências de blocos de anúncios sem considerar a sua envolvente programática, uma vez que não são emitidos em momentos fixos do dia e não têm personalidade própria, como os programas para as análises se poderem referir apenas as eles.

Concluindo, do ponto de vista da previsão dos valores de *ratings* programáticos os métodos quantitativos são essenciais, uma vez que as análises devem ser realizadas de

forma muito regular e para uma elevada dimensão de suportes, que devem abranger o horizonte temporal associado a uma campanha publicitária. Atendendo à elevada dimensão da base de dados, as análises devem ser realizadas da forma mais rápida e menos onerosa e com a menor intervenção dos planificadores. As previsões por juízos de valor apresentam inúmeras limitações que se encontram descritas na vasta bibliografia existente, entre as quais um sobreajustamento aos dados que também tende a ocorrer com algumas abordagens metodológicas, nomeadamente com as análises lineares de Regressão. Nas previsões qualitativas ocorre, frequentemente, a identificação de padrões na evolução dos dados que não são reais e resultam do ruído existente. Verifica-se uma limitada capacidade de processamento dos dados e uma tendência para simplificarem demasiado as questões. Apesar de poderem ter em consideração determinados factores mais qualitativos, nem sempre estes são interpretados objectivamente. As análises que dependem da experiência e prévio conhecimento dos utilizadores produzem resultados que não são generalizáveis.

Consequentemente, foram avaliadas as potencialidade de determinadas abordagens lineares e não lineares e globais ou locais. Verificámos que as análises de Regressão linear múltipla não apresentam uma satisfatória capacidade predictiva, uma vez que se verifica uma ausência de lineariedade das relações observadas e as variáveis com maior capacidade explicativa não são previsíveis com reduzidas margens de erro ou conhecidas. Porém, as análises simbólicas superam os resultados das análises mais tradicionais.

Do ponto de vista linear a melhor opção será por associação com $Tvr\%$, tendo-se consequentemente, também, realizado algumas análises de previsão a este indicador com resultados que também não são muito satisfatórios. As análises de Regressão linear tendem a um sobreajustamento aos dados e implicam que sejam realizados juízos de valor com as limitações que lhe são inerentes em algumas etapas deste processo, como, por exemplo, quando se avaliam os pressupostos que condicionam a sua aplicação.

As relações históricas identificadas podem não se observar no futuro e não se adequam aos contextos em que se observam observações atípicas.

As análises não lineares permitem em parte ultrapassar algumas destas limitações, por exemplo, através de uma maior compreensão das interacções entre as variáveis e em princípio as relações podem se manter válidas durante um maior horizonte temporal. As análises não lineares por aplicação das Árvores implicam uma menor intervenção do utilizador e interpretam o conjunto de dados produzindo conhecimento fundamentado. As Redes Neurais apresentam um funcionamento mais opaco e poderá se verificar um sobreajustamento aos dados, não existindo uma formulação teórica que defina um intervalo de confiança como na Regressão linear. Segundo *Danaher et al. (2011)*:

The bayesian forecasting methods perform better than previous methods, but a simple Regression model with program effects also produces forecasts with similar accuracy to that of the bayesian methods. ... Substantial gains in forecast accuracy are unlikely to eventuate, even from highly complex models. (p.2)

Danaher et al. (2011) referem que, nem sempre, uma maior complexidade das metodologias conduz a melhores resultados em análises de previsão e que dificilmente vão ser identificadas alternativas que produzam melhores resultados. Na nossa perspectiva, concordamos que devem ser avaliadas alternativas de análise que viabilizem a construção dos planos de *media* e nos seguintes capítulos descrevem-se determinadas propostas. A previsão dos valores dos *ratings* programáticos em alguns contextos poderá não conduzir a valores satisfatórios independentemente das metodologias que forem adoptadas.

CAPÍTULO VI

APRESENTAÇÃO DE PROPOSTA PARA O PLANEAMENTO DE *MEDIA*

Neste capítulo apresentamos algumas perspectivas de análise, que podem contribuir para que melhores resultados possam ser alcançados quando se pretende a construção dos modelos de exposição. Atendendo ao significativo número de contextos identificados e aos inúmeros recursos que, por vezes, as análises envolvem, determinadas propostas essencialmente relacionadas com a análise dos dados que constam dos ficheiros base de audimetria, não foram implementadas na prática. Porém, a totalidade das abordagens metodológicas que são descritas, vão ao encontro de determinadas conclusões substântivas que constam da bibliografia internacional da especialidade, ao nível das características que as análises de indução apresentam nas Ciências Sociais, assim como dos resultados das análises que foram realizadas aos “*ratings*” e que caracterizam os comportamentos de consumo televisivo do painel nacional. Avaliaram-se algumas das potencialidades de área metodológicas mais recentes e adequadas para a análise dos dados que constam das bases de dados de audimetria, e, também, foi considerada a repercussão de eventuais alterações futuras no panorama audiovisual nacional, ao nível da evolução destes dados.

VI.1. INTRODUÇÃO

As Ciências Exactas geralmente conseguem proporcionar resultados, que apresentam um elevado grau de precisão a respeito dos sistemas em análise, mas nas Ciências Sociais as previsões são classificadas como inexactas, direccionais e probabilísticas.

As Ciências Humanas recebem a designação de ciências não duras pela ideia da realidade que oferecem não ser determinista e não conseguirem oferecer leis tão objectivas. Apesar do limitado acesso da maior parte do painel à actual oferta televisiva, na presente tese, foi possível concluir que impera a variabilidade do consumo televisivo, e, frequentemente, também dos indicadores que dele resultam. Consequentemente, não será viável desenvolver modelos que proporcionem resultados com uma elevada precisão. As Ciências Humanas, por oposição às Ciências Exactas, apresentam um carácter muito menos preciso, limitando-se, frequentemente, apenas a generalizações e classificações de conduta.

Ao contrário do que se verifica em alguns países europeus, a *Televisão Digital Terrestre (TDT)* nacional apresenta, apenas, 5 ou 4 canais confirmados. Consequentemente, por agora, não se perspectivam grandes alterações no sector televisivo decorrentes da introdução da TV digital. Segundo algumas notícias que surgiram na comunicação social, em Portugal, a maioria das pessoas não vê grandes vantagens em aderir à TDT, uma vez que a aquisição das “boxes” ou de televisores compatíveis representa um custo e sem acrescentar grandes novidades em relação aos conteúdos. No início, falava-se de um 5º canal, posteriormente, de um canal em Alta Definição para se adicionar aos restantes, mas a única certeza, por agora, é a disponibilização dos quatro canais em sinal aberto. No futuro, com o acesso à TV digital, os telespectadores podem, eventualmente, vir a ter acesso a um maior número de canais, o que poderá no decorrer do tempo contribuir para uma maior variabilidade dos indicadores de audimetria, se alguns segmentos de telespectadores diversificarem o conjunto de programas e canais a que costumam assistir. Pelo contrário, outros segmentos podem-se fidelizar a um número mais restrito de canais, que em alguns casos podem inclusive coincidir, maioritariamente, com as opções televisivas que apresentavam anteriormente. Em alguns países Europeus, tem-se verificado uma tendência para a polarização e especialização das audiências.

Um segundo factor, a ter em atenção, será a alteração da tecnologia de medição das audiências televisivas. Até 2012, as audiências eram medidas pela empresa Marktest e recorrendo a uma tecnologia de audimetria activa.

Porém, a partir do início de 2013 e por um período ainda não conhecido, atendendo à contestação manifestada por algumas das estações televisivas, a medição das audiências será realizada pela empresa GFK e por aplicação da tecnologia “*audio-matching*”.

As características dos dados disponibilizados pela empresa GFK são muito semelhantes aos anteriores, apresentando uma caracterização do consumo televisivo ao segundo e a descrição das características sociodemográficas do painel. Nas primeiras semanas de utilização deste novo processo, os resultados apresentados diferem significativamente dos valores da Marktest, uma vez que a constituição do painel também, por exemplo, ao nível da percentagem dos telespectadores que têm acesso à TV Cabo. Tem-se verificado uma atitude de rejeição perante este novo sistema, por parte de algumas estações televisivas e agências de meios. Apesar das críticas, anteriormente, direccionadas à empresa Marktest, com esta nova tecnologia verificou-se uma forte alteração dos valores dos índices de audiência. Tem-se verificado uma penalização das audiências do canal TVI e da RTP1, que virá a ser privatizado ou concessionada a sua gestão, e uma subida nas audiências da SIC e dos canais por TV Cabo.

A TV Cabo tem sido apresentada como um concorrente, cada vez mais forte da televisão generalista, o que no actual contexto macroeconómico poderá suscitar grandes dúvidas. As mudanças que ainda podem ocorrer neste sector, podem vir a contribuir para uma maior variabilidade dos dados, mas, eventualmente, apenas no longo prazo, ou pelo menos quando a situação macroeconómica permitir à população um maior investimento na sua oferta televisiva. O Canal 2 poderá vir a ser extinto, mas, até ao momento, o seu desempenho também não o tornava atractivo do ponto de vista do planeamento de *media* e do investimento publicitário. O futuro da RTP1 permanece incerto ao nível das características da sua programação.

Por conseguinte, também são necessárias abordagens que viabilizem a construção dos planos de *media* em contextos de uma maior variabilidade dos valores de “*ratings*” e que, também, constam das bases de dados construídas através da tecnologia de audimetria da Marktest.

Independentemente do grau de exactidão dos dados, o que se pretende, na presente tese, é avaliar em que medida, e quais os instrumentos metodológicos que podem contribuir para uma construção mais eficiente dos planos, em função dos recursos disponíveis e mais, facilmente, acessíveis às agências de meios. Nem sempre, as necessidades de informação das estações televisivas, que utilizam estes dados para planificar as suas grelhas, coincidem com as necessidades das agências de meios, que os utilizam para construir os planos de *media*.

Os métodos alternativos de medição das audiências televisivas não podem substituir as análises de audimetria ou o *audio-matching*. Em inúmeros contextos, a utilização das sondagens de intenção de ver, como método de previsão que recolhe os interesses dos telespectadores e a intenção de seleccionar determinados conteúdos, não produz resultados fiáveis, assim como outras abordagens qualitativas, para além de não viabilizar o cálculo de indicadores que são essenciais à construção dos planos de *media*. Raramente o que é declarado corresponde ao que é efectivamente seleccionado, ainda que considerando apenas o curto prazo, como consequência de uma ocultação, deturpação e até de um real desconhecimento das futuras selecções programáticas e das decisões de ver TV.

Também ficou demonstrado, os inúmeros contextos de exposição não intencional e, consequentemente, não planificada *a priori* que podem surgir, assim como as situações em que os factores estruturais contribuem para uma ausência de exposição, também, não intencional. Uma vez que, actualmente, poderá não se verificar maioritariamente uma suposta regularidade na exposição televisiva, e que ficou confirmada com base nas análises de duplicação das audiências que foram realizadas, ficamos sem saber que informação realmente proporcionam os trabalhos tradicionais que recorrem aos inquéritos como instrumento de quantificação das audiências televisivas e de construção dos planos de *media*.

Os inquéritos não viabilizam uma recolha de dados realmente objectivos, e, também, inúmera informação que é essencial para a actividade de planeamento de *media*.

Apesar das limitações que o processo de audimetria, por vezes, poderá apresentar, a sua utilização é necessária, devendo-se, sim, continuamente aperfeiçoar esta tecnologia ou qualquer outra que a possa substituir no futuro.

Consequentemente, é necessário desenvolver abordagens de previsão quantitativas, que viabilizem a construção dos planos de *media* e alicerçadas nos sistemas de informação existentes. Apesar de escassas existirem, na bibliografia da especialidade, algumas propostas alternativas. A última questão em avaliação corresponde à que consta do seguinte parágrafo:

Questão 4: “Em função das análises explicativas e de previsão realizadas, quais as abordagens metodológicas e os recursos alternativos que podem permitir ao sector televisivo alcançar melhores resultados quando se pretende desenvolver um plano de *media*, nos contextos de maior variabilidade dos dados, uma vez que a maior parte das receitas publicitárias decorre das vendas realizadas no “up-front market”?”

Na presente tese, avaliam-se eventuais potencialidades da Análise Simbólica e das diversas vertentes de análise ligadas ao *Data Mining*, que corresponde a uma área muito abrangente enquadrada na Engenharia Informática. Privilegia-se a avaliação de determinadas áreas de DM não proposicional, como, por exemplo, as áreas de DM Sequencial, Temporal e Multirrelacional. O DM não proposicional viabiliza a análises de bases de dados que estão caracterizadas por uma representação mais rica da informação, como, por exemplo, considerando as bases de dados relacionais ou uma sequência de selecções realizada por determinado telespectador ou conjunto de telespectadores. A Análise Simbólica permite considerar variáveis que caracterizam, de forma mais adequada, os indicadores que apresentam uma considerável variabilidade.

Pretende-se avaliar, essencialmente, os modelos de exposição, apesar da abordagem sugerida permitir, simultaneamente, a selecção dos suportes que cumprem determinados requisitos ao nível do perfil das suas audiências ou da cobertura de exposição, e, assim, parcialmente “optimizar” a selecção dos mesmos.

As restrições orçamentais, entre outras, devem fazer parte, posteriormente, de algoritmos de optimização que não são descritos exaustivamente na sequência metodológica que é sugerida.

A informação necessária em determinadas etapas do processo que permite construir os planos de *media*, poderá transcender a que consta das bases de audimetria. Por vezes, deverá ser complementada com outras fontes primárias de dados e, tendo em atenção as diversas perspectivas explicativas da exposição televisiva, destacando-se como provavelmente mais relevante a perspectiva Estrutural.

Devemos, também, considerar o comportamento agregado, uma vez que corresponde ao que realmente condiciona a venda do espaço comercial. A informação, que consta dos ficheiros base de audimetria, tem que ser agregada para ser conhecida a percentagem de audiência que assiste a determinado programa e com que regularidade. Se um telespectador estiver exposto a determinado suporte poderá não ser equivalente a “ver” o mesmo. Contudo, estes dados são necessários, uma vez que representam uma possibilidade de exposição a uma inserção publicitária.

As bases de audimetria, posteriormente, devem ser utilizadas para prever as audiências televisivas. Será necessário medir, com uma cada vez maior precisão e exactidão, a exposição e a relação dos indivíduos com cada programa, o que a audimetria certamente viabiliza comparativamente aos métodos alternativos de medição das audiências. Uma vez que nestas bases de dados consta uma escassa informação de caracterização dos conteúdos transmitidos, assim como do painel será necessário complementar estes dados com outras fontes que permitam ultrapassar determinadas lacunas informativas.

A construção de um plano abrange várias etapas que se diferenciam quanto ao grau da necessidade de intervenção do planificador. Inicialmente, será necessário definir o público objectivo que deverá ser seleccionado, em função de variáveis sociodemográficas e psicográficas. Porém, o painel nacional apenas é caracterizado por variáveis sociodemográficas.

Uma subamostra poderá ser seleccionada, inicialmente, em função das especificações do público-alvo da campanha. Recorrendo a análises de segmentação podem ser identificadas classes de indivíduos, em função dos seus perfis de exposição televisiva, e de entre os quais são seleccionados os que mais se ajustam ao público-alvo de um anunciante. Para cada segmento em contextos de maior variabilidade dos dados devemos identificar as sequências de selecções mais características. Identificados os conteúdos a que os vários segmentos se encontram expostos em maior grau, independentemente das manifestações mais concretas que se verificam em cada dia e semana, a selecção dos suportes considerados como mais adequados poderá ser complementada por análise das associações existentes entre os suportes e as variáveis de caracterização do painel ou face aos públicos objectivos.

Alguns telespectadores com comportamentos diferenciados em determinado dia e semana podem corresponder a um mesmo segmento, uma vez que se pretende a análise da exposição global ao *media* e da estrutura dos seus comportamentos. A frequência de exposição que decorre da construção das Regras de Associação e/ou Sequenciais será a que lhes é mais característica, e como tal, eventualmente, mais eficaz, uma vez que se consideram os conteúdos que na realidade se complementam e se enquadram no conjunto de conteúdos que cada telespectador poderá vir a seleccionar em maior grau no curto prazo.

Os resultados devem ser comparados com os resultados de planos que são construídos manualmente. Neste processo os sistemas informáticos são insuficientes, a experiência, conhecimentos e intuição de um planificador é determinante (*Makienko, 2012*). Para conseguir previsões mais concretas para determinado período será necessário recorrer aos modelos explicativos da exposição televisiva e a inúmeras metodologias de *Data Mining* que viabilizem a construção de Regras Relacionais, que permitem a avaliação das relações existentes entre as diferentes entidades representadas na base de dados.

Numa etapa prévia, apesar de se considerar as exposições que são naturais a cada telespectador e segmento e não as “frequências efectivas” que corresponde a um conceito que levanta inúmeras críticas e é considerado como pouco fiável (*Jones, 1995*), os suportes a seleccionar podem ser posteriormente ajustados por forma a aumentar ou diminuir a intensidade dos contactos em função das características da campanha.

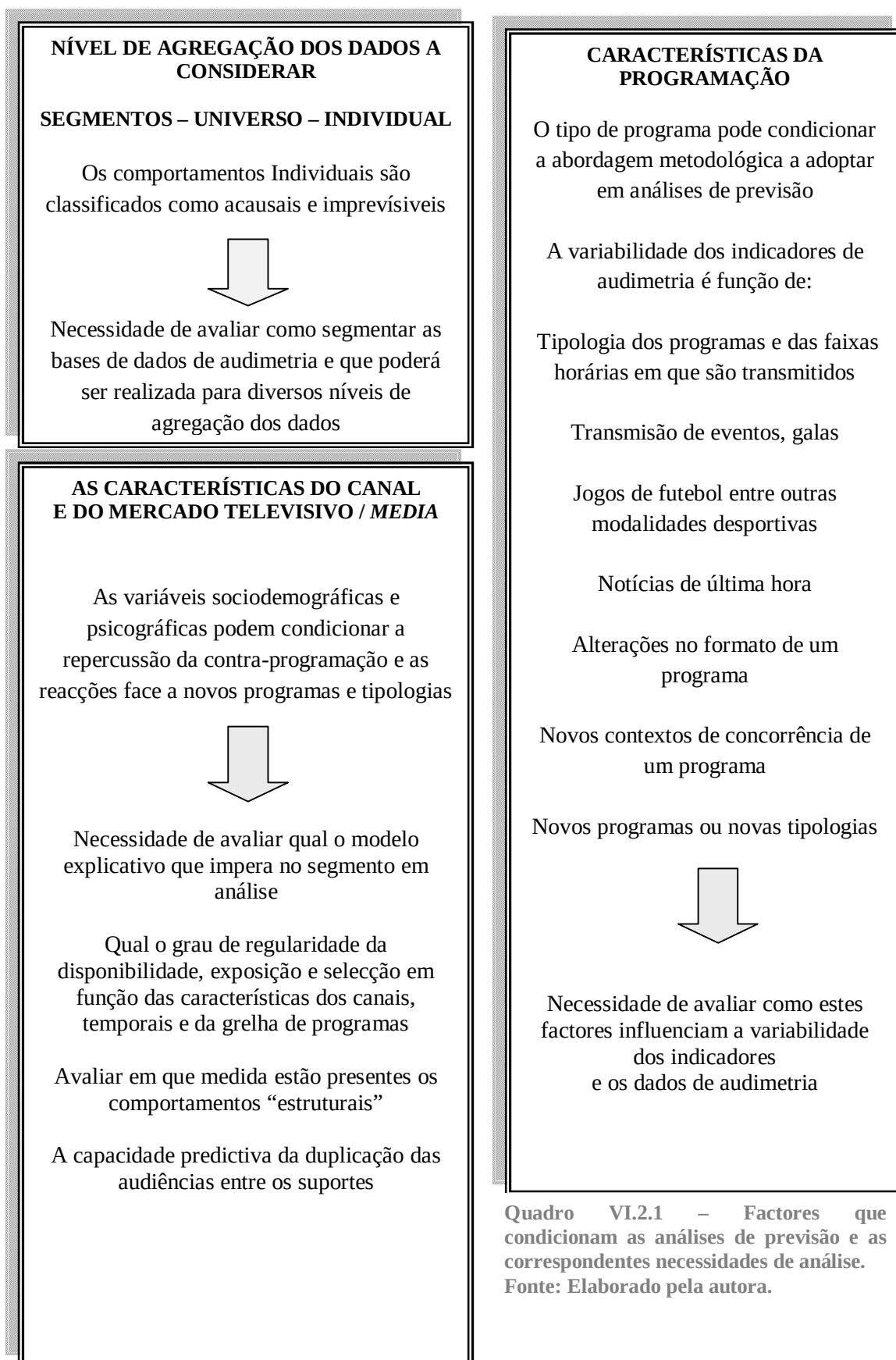
Definido um plano que seja considerado como o mais adequado para comunicar com determinado segmento será necessário construir um solução final que, também, será função do orçamento disponível e que requer a aplicação de algoritmos de optimização e ter em consideração os resultados das análises realizadas nas etapas anteriores. Uma vez que os comportamentos são dinâmicos e atendendo à sazonalidade que caracteriza a exposição televisiva, os resultados têm que ser reavaliados no decorrer do tempo, mas, tirando partido da informação e dos modelos históricos que foram construídos.

Nos seguintes capítulos, descreve-se em maior detalhe uma proposta para a construção dos planos de *media* e em função dos contextos identificados na evolução dos dados.

VI.2. ENQUADRAMENTO METODOLÓGICO DA PROPOSTA PARA PROMOVER A CONSTRUÇÃO DOS MODELOS DE EXPOSIÇÃO

A previsão das audiências televisivas corresponde a uma área muito abrangente, atendendo à diversidade de indicadores que constam das bases de audimetria e aos contextos que requerem a aplicação de diferenciadas perspectivas de análise.

O *Quadro VI.2.1* e o *Quadro VI.2.2* apenas ilustram alguns dos factores que condicionam em maior grau as abordagens de previsão a serem adoptadas. Os principais factores que influenciam as análises correspondem ao perfil do público-alvo, ao nível de agregação dos dados e indicadores em análise, aos factores temporais, às características dos programas e de estruturação da grelha, à contra programação e ao canal, assim como se nos referimos a indicadores programáticos, horários ou comerciais ou se estão em análise os comportamentos estruturais entre os suportes publicitários.

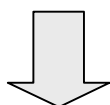


Quadro VI.2.1 – Factores que condicionam as análises de previsão e as correspondentes necessidades de análise.

Fonte: Elaborado pela autora.

CARACTERÍSTICAS DE ESTRUTURAÇÃO DA GRELHA E DA CONTRA PROGRAMAÇÃO

As estratégias de programação condicionam a evolução dos indicadores, por exemplo, o “Hammocking”, “Tent-Poling”, “Lead_In”, “Lead_Out”, entre outras

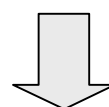


Necessidade de avaliar como a programação dos canais concorrentes, por mimetismo ou em contra-programação, condicionam a evolução dos indicadores

OS FACTORES TEMPORAIS

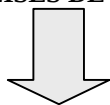
As faixas horárias, o mês, o dia, a semana versus fim de semana, os feriados e pontes condicionam a evolução dos indicadores

As metodologias são função do prazo a considerar na previsão: curto ou médio



Necessidade de avaliar a variabilidade que decorre das variáveis temporais e suas implicações a nível metodológico

PÚBLICO-OBJECTIVO DAS ANÁLISES DE PREVISÃO



Necessidade de avaliar qual o modelo explicativo que impera no segmento em análise

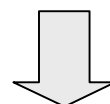
Qual o grau de regularidade da disponibilidade, exposição e selecção em função das características dos canais, temporais e da grelha de programas

Avaliar em que medida estão presentes os comportamentos “estruturais”

A capacidade predictiva da duplicação das audiências entre os suportes

PREVISÃO ASSOCIADA À COMPONENTE COMERCIAL E À INTER-RELAÇÃO ENTRE AS AUDIÊNCIAS COMERCIAIS E AS PROGRAMÁTICAS

As audiências comerciais dependem das audiências programáticas e apresentam uma evolução que é função das características dos espaços comerciais



Necessidade de avaliar em que medida os factores de planeamento comercial condicionam a evolução dos *ratings* comerciais ou do seu valor relativo face às audiências programáticas que lhe estão associadas

Quadro VI.2.2 – Factores que condicionam as análises de previsão e necessidades de análise.
Fonte: Elaborado pela autora.

O sector tem que satisfazer as necessidades dos anunciantes através da previsão de inúmeros indicadores, que por sua vez condicionam as perspectivas de análise a serem adoptadas, por exemplo, a avaliação da afinidade que define o grau de adequabilidade de determinada faixa horária para o segmento-alvo do anunciante; a análise da cobertura que quantifica o número de indivíduos expostos, pelo menos uma vez, a uma campanha; a frequência de exposição e os *ratings*, entre muitos outros indicadores.

Sendo as Ciências Sociais inexactas, e os comportamentos individuais classificados como acausais e imprevisíveis, devem ser combinadas diversas abordagens e desenvolvida uma teoria do erro, privilegiando-se a análise dos dados agregados. Não existe uma única perspectiva de análise e devemos em função dos contextos avaliar, no conjunto de abordagens realmente exequíveis, as que apresentam maiores vantagens. Os dados de audimetria contribuem para esclarecer inúmeras incógnitas relativamente ao consumo televisivo e constituem o instrumento de planificação e de previsão por excelência.

Neste capítulo, consta uma síntese das perspectivas de análise que podem contribuir para a construção dos modelos de exposição, recorrendo aos valores de *ratings* para o universo televisivo e/ou por segmentos, assim como quando é necessário considerar os fluxos de audiências entre os suportes publicitários (*Fig. VI.2.1*). Ainda que na bibliografia da especialidade seja privilegiada a análise dos *ratings* e em alternativa as metodologias qualitativas de previsão, os planos de *media*, também, podem ser construídos recorrendo a abordagens quantitativas alternativas e que constam do estado da arte apresentado.

Atendendo à escassa bibliografia que se encontra disponível na área da construção dos planos de *media* e dos modelos explicativos e de previsão dos “*ratings*”, inclusive a nível internacional, nas análises realizadas e sugeridas constam abordagens metodológicas, que pretendem ir além da Estatística Multivariada mais tradicional.

Do ponto de vista da aplicação do *Data Mining* são reduzidos os contributos que constam da bibliografia da especialidade, incidindo, essencialmente em algumas metodologias proposicionais e não foram identificadas Análises Simbólicas.

Inicialmente foi necessário avaliar a viabilidade de construção de modelos de *ratings*, que será certamente a perspectiva que exige uma menor intervenção do utilizador e mais rápida do ponto de vista da realização das análises e construção das bases de dados e com menores custos envolvidos. Atendendo aos requisitos que os sistemas de previsão apresentam, os modelos de “*ratings*” devem ser a perspectiva privilegiada. As análises podem incidir no universo televisivo ou em segmentos que devem ser previamente identificados.

Nas análises realizadas procurou-se a identificação dos segmentos que podem resultar das variáveis de segmentação das bases de dados de audimetria em função da variabilidade que apresentam os principais indicadores. O objectivo de um estudo de audiências corresponde a caracterizar e quantificar a exposição televisiva, sendo os indicadores agregados os que realmente são mais relevantes e que devem ser alvo de análises de previsão e/ou explicativas, uma vez que actuando sobre as variáveis controláveis podemos influenciar a sua evolução.

A exposição individual não é relevante por si só, mas, sim, a evolução de determinados indicadores de audimetria, assim como a exposição global do público-alvo face a um conjunto de suportes, que devem ser seleccionados em função de determinados critérios quantitativos e qualitativos. O conteúdo, nem sempre, é responsável por uma forte variabilidade da evolução dos indicadores de audimetria. A transição entre os canais nem sempre penaliza significativamente a evolução dos *ratings* de determinado suporte, sendo, frequentemente, mais relevantes as decisões de ligar e desligar o televisor que, também, podem ser independentes dos conteúdos e imperar a “experiência” televisiva. Uma maior percepção dos factores que condicionam a evolução dos dados é essencial para a construção das campanhas publicitárias.

Em Portugal, para o período em análise grande parte do painel não apresenta TV Cabo, e, logo, o número de opções televisivas alternativas não é tão abrangente como se menciona, por vezes, a nível internacional. Verificamos que para alguns segmentos existe alguma regularidade na exposição. Consequentemente, em determinados contextos, as análises de séries temporais podem produzir resultados satisfatórios atendendo à dimensão dos erros de previsão que, frequentemente, constam da bibliografia internacional da especialidade.

Nas análises teve-se em consideração como variável explicativa e de previsão, as tipologias programáticas, recorrendo ao nível mais desagregado de classificação dos conteúdos que constam das bases de dados, e considerando do ponto de vista simbólico as variáveis taxonómicas que permitem avaliar a repercussão desta variável a diferentes níveis de agregação dos dados.

Segundo a perspectiva explicativa da exposição televisiva que incide na análise das gratificações e usos do *media*, os telespectadores apresentam preferências por tipologias, e não, por programas em concreto, contudo, as preferências, por si só, não determinam a exposição. Por conseguinte, foi necessário desenvolver um modelo, essencialmente estrutural e explicativo da exposição, que deverá ser complementado com variáveis individuais e considerado nas análises de previsão dos fluxos de audiências entre os suportes.

Com base nas análises realizadas aos índices de duplicação verificamos que, a nível nacional, existe um segmento de telespectadores que apresenta alguma duplicação de audiências, inclusive considerando um maior desfasamento temporal entre os suportes seleccionados, podendo-se concluir que, neste contexto, existe alguma regularidade na exposição televisiva. Apesar da regularidade caracterizar o comportamento de um menor número de telespectadores, estes segmentos podem ser mais valorizados pelos anunciantes, e, logo, as melhores perspectivas de previsão devem ser identificadas, também, neste contexto, assim como quando impera a irregularidade dos padrões de consumo televisivo e dos *ratings*.

Nos contextos de maior irregularidade dos dados, uma vez que a globalidade dos programas a que se encontram expostos os telespectadores, também, deve ser considerada, aborda-se a implementação das Regras de Associação ou Sequenciais, inclusive ao nível do *Data Mining* Incremental em contextos mais dinâmicos, assim como a análise da duplicação de audiências entre os diferentes suportes publicitários.

Foram privilegiadas as análises que apresentam um cariz explicativo da evolução dos dados, uma vez que interessa a identificação da estrutura dos comportamentos que originam determinadas evoluções dos índices de audiência. O padrão de ver TV que deve ser identificado corresponde à estrutura dos comportamentos de consumo televisivo, que não necessariamente deriva numa relação de causalidade face aos interesses e motivações do painel. Nas análises teve-se em atenção privilegiar as potencialidades de áreas metodológicas que viabilizem a análise dos fluxos de audiências e das interdependências existentes, não sendo os dados a nível individual relevantes, por si só, e em especial se considerarmos um suporte em concreto.

Neste capítulo constam algumas sugestões que podem viabilizar a previsão de valores de *ratings* ou a construção dos modelos de exposição através de perspectivas alternativas de análise. Num contexto de maior estabilidade, e que incide, essencialmente, nas faixas horárias que antecedem o pré-prime time em dias laborais, ou ao fim de semana que em função das características da programação poderá abranger a totalidade da emissão diária, os melhores resultados podem ser viabilizados recorrendo a análises de séries temporais ou a análises multivariadas de Regressão de valores de *ratings*. Em contextos de maior variabilidade será necessário recorrer ao *Data Mining* não proposicional e complementar com outras fontes de dados. Nomeadamente devemos recorrer aos dados dos ficheiros de base de audimetria e a inquéritos que permitem, simultaneamente, avaliar e caracterizar em maior profundidade o que determina a exposição televisiva.

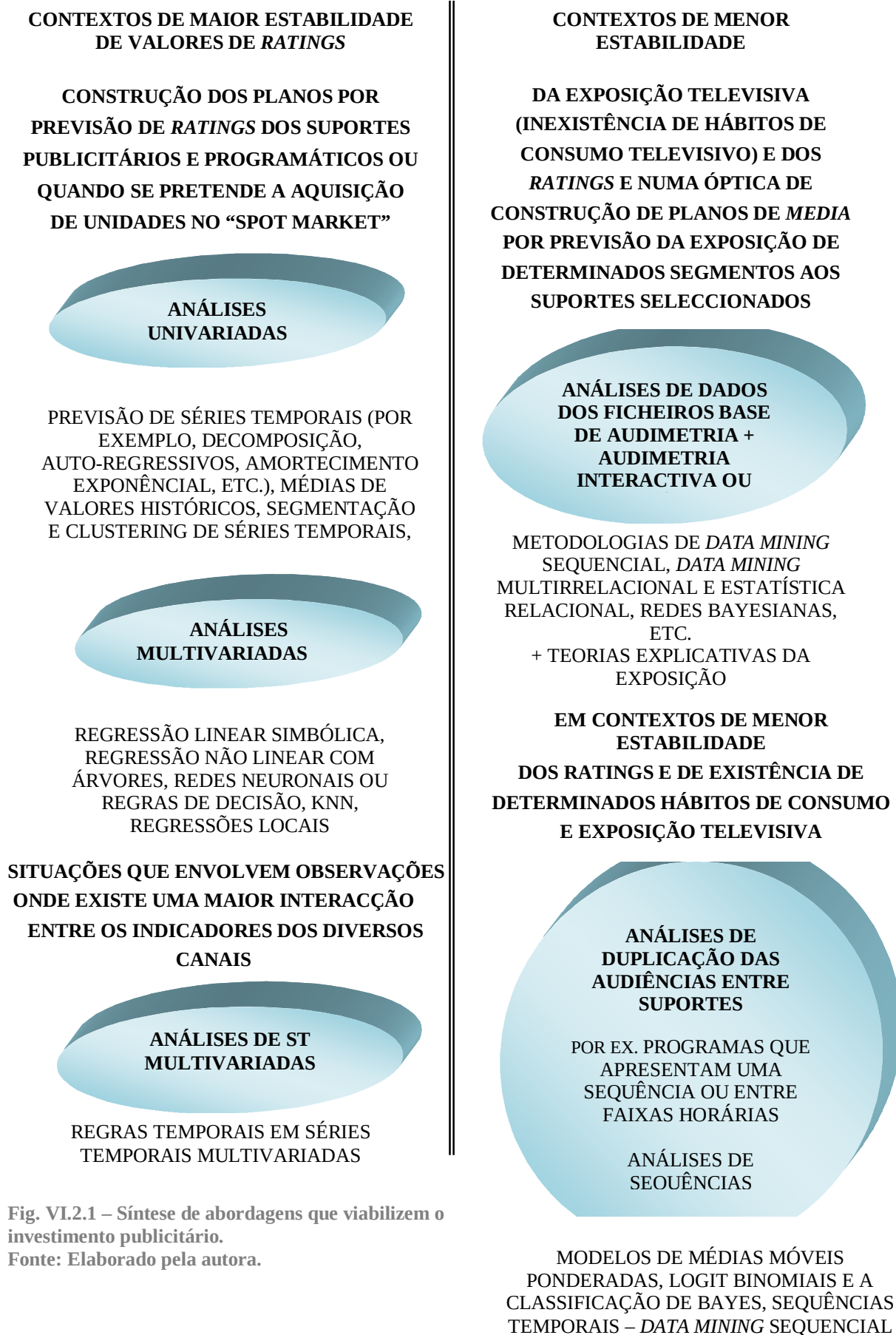


Fig. VI.2.1 – Síntese de abordagens que viabilizem o investimento publicitário.

Fonte: Elaborado pela autora.

VI.3. OS MODELOS DE EXPOSIÇÃO EM CONTEXTOS DE MAIOR ESTABILIDADE DE VALORES DE RATINGS

Atendendo aos resultados das análises realizadas, e que se diferenciam significativamente em inúmeros contextos do que se verifica a nível internacional, defendemos que a previsão de valores de *ratings* poderá ser uma opção a considerar e poderá ser realizada através de diversas abordagens, entre as quais *a priori* se destacam, quer as análises univariadas, quer as análises multivariadas.

VI.3.1. As análises univariadas de valores de *Ratings*

Verificamos que no decorrer de determinadas faixas horárias, que abrangem, não só, as mais estáveis podemos ignorar os restantes indicadores e factores estruturais e ter em atenção apenas a sua evolução histórica. Existem contextos em que uma mera extrapolação de valores passados, ignorando a constituição da grelha constitui uma opção, nomeadamente a utilização das metodologias simbólicas de previsão (*Maia, Carvalho & Ludemir, 2006*) ou a segmentação (*Lemire, 2006*) e clustering (*Prakash & Li, 2011*) das séries temporais.

A nível prático, frequentemente, para a totalidade da emissão diária, os valores de *ratings* são calculados apenas em função dos seus valores históricos para alguns escassos anos e considerando idênticos períodos horários. Quando o horizonte temporal da campanha publicitária é mais abrangente, em virtude da estacionalidade que caracteriza o consumo televisivo, será necessário o cálculo de coeficientes mensais, por faixa horária e por público-objectivo. As metodologias mais adequadas procedem, inevitavelmente, do *Data Mining* Temporal (*Laxman & Sastry, 2006*) e mais concretamente do *Data Mining* de séries temporais (*Esling & Agon, 2012*).

Em apêndice, constam algumas análises, meramente exploratórias, com o intuito de tentar prever com análises univariadas os valores de *ratings* e de taxas médias de audiências para o total da TV, recorrendo a metodologias simbólicas e de *Data Mining*, ao nível da segmentação e *clustering* das séries temporais.

As análises, por vezes, conduziram a bons resultados, em especial se considerarmos os valores dos indicadores para as faixas horárias. Considerando as respectivas séries temporais para menores unidades de análise, como, por exemplo, ao segundo, e por via da segmentação e *clustering* melhores resultados, certamente, podem ainda vir a ser alcançados, apresentando-se esta, frequentemente, como a melhor opção, atendendo ao conjunto de metodologias que forma avaliadas.

A nível nacional e para o horizonte temporal em análise, não se verifica a inadequabilidade destas análises, frequentemente, referida, actualmente, na bibliografia internacional. Contudo, os trabalhos publicados referem-se, essencialmente, à aplicação dos algoritmos mais tradicionais de previsão, por exemplo, a análises ARIMA (Makridakis, Wheelwright & Hyndman, 1997) ou Holt Winters (Tavares et al., 1996), entre outras, por análise de tendências e da sazonalidade, que na realidade podem não ser os mais adequados. Na bibliografia da especialidade, frequentemente, associam à previsão quantitativa apenas as análises que não consideram variáveis explicativas da evolução das séries temporais e procuram apenas a sua tendência e variações estacionais. Mas, na realidade as análises que incidem nas séries temporais multivariadas podem ir muito mais além. Apesar de existirem métodos de previsão quantitativos mais elaborados, como os *Box-Jenkins* são poucos os autores que recorrem a estas abordagens (Danaher et al., 2011).

Apesar da essência da evolução de uma série temporal não se encontrar na simples enumeração de uma série histórica de observações, se o objectivo das análises residir apenas na previsão dos seus valores para o curto prazo, em alguns contextos a análise das ST, segundo abordagens menos tradicionais, deverá ser uma opção a considerar. Na opinião de Bunge (1981), as linhas de tendência como instrumentos de previsão não são dignas de confiança se não se basearem em teorias que expliquem os mecanismos que causam essa tendências que estão na sua base. Salienta que nas análises de interpolação, também, devemos preferir a simplicidade formal e matemática, uma vez que nada nos garante que seja melhor adoptar um modelo mais complexo. O autor, também, salienta que o passado remoto não é relevante para as análises de previsão, uma vez que se podem produzir significativas alterações no cenário dos *media*.

A nível nacional, verificamos que, nem sempre, no decorrer dos anos a variabilidade observada nos dados é superior face à que existe no decorrer de alguns meses, apesar de se poderem verificar algumas alterações no conteúdo da grelha de programas e ao nível do desempenho global dos canais. Além do mais, também, foi considerado nas análises realizadas um horizonte temporal que apresenta significativas alterações no desempenho global das estações televisivas, e, logo, poderia ser mais variável. Nas faixas horárias que antecedem o horário nobre, entre outras onde existe uma maior estabilidade dos indicadores, em especial considerando o curto prazo, podemos projectar com a devida ponderação os dados históricos. Assim, podem ser realizadas previsões, a não ser que os canais realizem importantes alterações nas suas grelhas de programas.

Frequentemente, de ano para ano, apesar das alterações existentes nos programas que integram a grelha, as tipologias transmitidas são mantidas em faixas horárias equivalentes. A influência das alterações realizadas nas grelhas, nem sempre, condiciona fortemente os resultados, uma vez que as estações televisivas optam por manter as tipologias transmitidas. Por exemplo, na grelha de 2010-2011, a TVI substituiu um concurso ao final da tarde (“Quem Quer ganha”) por outro programa que apresenta características semelhantes e, posteriormente, retomou o formato original. O canal estatal RTP1, em 2012, transmite no período da noite um concurso que veio substituir um formato anterior da mesma tipologia programática.

Os canais generalistas nas faixas horárias da noite recorrem sistematicamente à transmissão de telenovelas, desde há longa data. A SIC, a partir de determinada altura, começou a transmitir as telenovelas brasileiras ao início da noite e apostou, também, na ficção nacional, logo após o espaço informativo da noite. Excepcionalmente, em parte do ano de 2012, a segunda telenovela da noite, também, foi produzida pela Rede Globo, assim como após o verão de 2013. Ao fim de semana, a TVI, aposta ao Domingo à noite nos “reality shows” ou em concursos, nomeadamente em programas familiares de entretenimento, tendo a SIC também adoptado, por vezes, este formato com a transmissão de concursos de cariz musical, entre outros.

A estabilidade que, frequentemente, observamos nas tipologias para a totalidade dos canais deriva da estratégia de cada uma das estações televisivas, e que procura maximizar o respectivo valor de SHR%.

Foram realizadas algumas análises de previsão não linear a *Tvr%*, e verificamos que inclusive os algoritmos de Regressão não linear do *Memory Based Learning* (MBL) (Cunningham & Delaney, 2007) e as Redes Neurais MLP (Chow, 2007; Khoshhal & Mokaram, 2012; Tang, 2007) nem sempre produzem resultados muito satisfatórios. O valor do erro quadrático médio relativo apresenta o valor máximo de 75,7%, considerando os dados dos dias laborais e recorrendo a Regressões Locais. A Rede MLP da tipologia 7-4-1 apenas apresenta 57,6% de observações previstas correctamente classificadas nos intervalos originais, se considerarmos intervalos de 3% de amplitude de valores de *Tvr%*.

Segundo as análises de segmentação realizadas a *Tvr%*, recorrendo ao *clustering* divisivo e simbólico dos dados e incidindo num período característico a nível médio do consumo televisivo podemos alcançar os 90% de variância explicada para uma solução a cinco *clusters* que apresentam entre si um diferencial de valores entre os 3,5% aos 5% e de 10% se consideramos na amostra as observações atípicas. Nas análises anuais, apesar das fronteiras entre os agrupamentos não se diferenciarem de forma muito significativa, para os diversos anos considerados, a sua constituição difere aos nível dos períodos horários que abrangem. Por vezes, os diferenciais entre cada agrupamento são reduzidos, mas, atendendo ao reduzido a moderado valor deste indicador, a nível relativo, a variabilidade pode ser considerada significativa.

Os pontos de inflexão na evolução dos dados encontram-se temporalmente mais distanciados para os dias laborais. Considerando a base de dados global, identificamos oito agrupamentos, como sendo os mais representativos da evolução desta variável no decorrer da emissão diária. Verificamos que períodos não sequenciais podem apresentar um comportamento comparável, sendo as audiências superiores e mais variáveis, entre as 20H e as 23H, e em especial em dias semanais.

Considerando os valores de *ratings*, as Regressões não lineares, também, apresentam potencial de aplicação em determinados contextos, nomeadamente as Regressões Polinomiais, Trigonométricas e as Regressões Locais, assim como o *clustering* das séries e considerando neste último contexto os protótipos de cada um dos agrupamentos como o valor da previsão. Porém, estas abordagens têm que incidir nos contextos de maior estabilidade dos dados, sendo o factor temporal uma das variáveis que os permite diferenciar, mas não só. Os valores de *ratings* calculados para os períodos horários são mais estáveis face aos correspondentes valores programáticos.

Através do *clustering* das séries temporais verificamos que, em função do canal, existem evoluções que podem ser consideradas representativas para a totalidade da semana ou diferenciando o período semanal e de fim-de-semana. Os dados dos canais públicos são mais estáveis, e a variabilidade dos indicadores é uma função do desempenho global do canal. Os restantes agrupamentos incluem as evoluções mais atípicas, que decorrem essencialmente da transmissão de alguns eventos desportivos.

Por conseguinte, os valores programáticos apresentam uma maior variabilidade, como consequência das suas oscilações horárias ou ao nível das suas durações, assim como devido ao grau de permanência das audiências no seu decorrer e que será função da exposição global ao *media*. Uma vez que a exposição global ao *media*, quantificada através da variável *Tvr%*, é mais estável em períodos de lazer, deverá ser superior a influência do grau de actividade do seu público na evolução dos *ratings*. Os *ratings* representam a percentagem de telespectadores presentes, em média, num minuto do suporte considerado nas análises.

Verificamos que existe alguma independência da evolução dos *ratings* face às alterações verificadas na construção das grelhas de programas, se considerarmos os dados referentes a períodos horários.

Por exemplo, por aplicação de algoritmos de *clustering* simbólicos aos dados de *ratings* das faixas horárias do canal RTP1, em alguns anos e para uma solução com quatro agrupamentos podemos alcançar os 97% de variância explicada, o que corresponde a uma variabilidade de *ratings* com o valor máximo de 5% para a maior parte dos agrupamentos considerados.

Nas análises em que se recorreu às Redes Neurais SOM simbólicas (Yang, Hung & Chen, 2012) e à amostra que integra a totalidade dos anos em análise, se consideramos apenas os dados do canal SIC e para uma solução que inclui quatro agrupamentos, o grau da variabilidade dos valores deste indicador encontram-se no intervalo 3,4% a 14,3%, sendo estes valores de 4,5% a 15% para a TVI e de 4,7% a 12,3% para a RTP1, desde que excluindo as observações atípicas. A rotatividade do painel, quer dizer, as suas substituições parciais no decorrer do tempo, também, poderá contribuir para estes valores. As abordagens multivariadas, entre outras perspectivas de análise, devem ser avaliadas nos contextos de maior variabilidade dos dados.

VI.3.2. As análises multivariadas de valores de *ratings*

Nas análises multivariadas diversas abordagens podem ser adoptadas, nomeadamente por avaliação da interacção observada entre as séries temporais dos três principais indicadores de audimetria; por análise da associação de valores de *ratings* e de taxas médias para o total da TV ou recorrendo a análises multivariadas de Regressão linear ou de Regressão não linear e considerando um conjunto de variáveis explicativas mais abrangente e que devem ser conhecidas no momento de realizar a previsão.

VI.3.2.1. A análise de séries temporais multivariadas

Em determinados contextos podemos considerar as interdependências existentes entre os três principais indicadores, ou seja, os *ratings*, as taxas médias para o total da TV e os *shares* de audiências.

A nível nacional, aparentemente esta abordagem do ponto de vista da previsão deve ser, essencialmente, considerada em contextos de transmissão de programas que geram observações atípicas, por exemplo, os jogos de futebol, entre outras transmissões desportivas.

O acréscimo que, por vezes, se verifica nas taxas médias de audiência para o total da TV, como consequência da transmissão de programas de maior popularidade, repercute-se de forma muito diferenciada nos restantes indicadores, podendo, por vezes, influenciar apenas um canal, e, não, se repercutindo nos canais concorrentes, ou quando se verifica variando consideravelmente na sua dimensão e extensão. Porém, uma vez que a curto prazo não existe um significativo número de situações que possam ser consideradas muito semelhantes à observação em análise, será necessário considerar um maior horizonte temporal.

Nas situações em que se opte por esta perspectiva de análise, do ponto de vista metodológico, será necessário considerar determinadas metodologias de *Data Mining* Temporal, nomeadamente as Regras Temporais (Guimaraes, 2000), que permitem avaliar as interações que se verificam em séries temporais multivariadas. Estas metodologias identificam os cenários históricos, ou seja, quais são as alterações que se verificam nos indicadores em estudo, por exemplo, acréscimo ou decréscimo em determinados intervalos temporais coincidentes ou próximos.

Os valores da previsão devem corresponder à média dos valores históricos observados em contextos mais “equiparáveis”, ou recorrer a variáveis simbólicas que caracterizem a variabilidade observada. As abordagens qualitativas permitem, também, tentar definir coeficientes correctores dos níveis de audiências, que podem ser alcançados nos contextos em que não é transmitida esta tipologia de programas. Nestes contextos será relevante dispor de informação que interprete a direcção da evolução mais provável dos dados, independentemente de ser viável uma quantificação mais precisa da sua variabilidade, e que, por conseguinte, permita aos anunciantes realizar uma melhor negociação com a estação televisiva em função das previsões direccionais e das

tendências mais prováveis de serem observadas nos dados. A previsão poderá corresponder apenas a uma função dos valores de *ratings* históricos ou resultar de um modelo em que se associe esta variável aos valores de *Tvr%*. Os modelos causais, dificilmente, podem ser construídos para esta tipologia de programas.

Recorrer a estes três indicadores sob outras perspectivas será muito problemático, uma vez que os *shares* de audiência colocam grandes obstáculos às análises de previsão. Na opinião de *Cattin et al. (1994)*, é extremamente difícil conseguir prever os *shares*, sendo necessário combinar vários modelos. A previsão poderia ser conseguida por extrapolações simples e, logo, ignorando o conteúdo das grelhas de programação ou por modelos de participação que possam distribuir a audiência total por cada canal, porém, *Rust e Alpert (1984)* apresentam um modelo que descrevem como sendo um protótipo que tem que ser significativamente melhorado. Os autores, calculam os valores de *shares* por identificação dos cenários históricos mais semelhantes à situação a prever, e quando o número de observações é reduzido tem que se recorrer aos seus valores médios. Os erros de previsão podem ultrapassar os 10% e as análises incidem num horizonte temporal que apresenta uma maior estabilidade face ao que se verifica actualmente. Todavia, para as faixas horárias que decorrem entre as 22H e as 23H apresenta erros com uma ordem de grandeza de 29,7% e entre as 23H e as 24H com um valor de 24,2%.

Com a aplicação de metodologias multivariadas de Análise Simbólica e de *Data Mining* conseguimos, a nível nacional, reduzir esta dimensão dos erros de previsão. Em *Cattin et al. (1994)*, também, é salientado que qualquer modelo a ser construído deve ser comparado com os resultados de análises mais simplificadas, para avaliar se se justifica o seu desenvolvimento. Com base nas análises realizadas, também, verificámos que os valores de *shares* apresentam uma variabilidade que transcende a que foi observada para os *ratings*, não constituindo *a priori* uma variável com uma razoável capacidade explicativa ou de previsão. Porém, para a identificação de cenários e para a construção de Regras Temporais a sua utilização poderá ser relevante.

Em alguns contextos, apesar de se verificar uma significativa redistribuição ao nível dos valores de *shares*, os valores de *ratings* podem apresentar uma maior estabilidade, a qual será função do grau da variabilidade da exposição global ao *media*. Nestes contextos, a maior condicionante será a viabilidade de identificação de um significativo número de cenários mais semelhantes à observação a prever.

Verificámos que existem vários cenários possíveis e uma grande diversidade ao nível da repercussão dos eventos desportivos, em função da relevância do jogo, equipas envolvidas, faixas horárias de transmissão, canal, entre muitos outros factores. Porém, em *Nikolopoulos et al. (2006)*, os autores analisam uma amostra de dados de audimetria gregos e conseguem alcançar resultados com alguma precisão, apesar do reduzido número de observações que foram utilizadas nas análises.

Em suma, a análise das séries temporais multivariadas dos principais indicadores de audimetria devem ser utilizadas para identificar tendências nos dados, permitindo aos anunciantes ou aos seus representantes realizar uma melhor negociação com as estações televisivas. Algumas metodologias de *Data Mining* Temporal, por exemplo, ao nível das Regras Temporais, viabilizam previsões direccionais.

VI.3.2.2. A associação de *ratings* e de taxas médias de audiências para o total da TV

Em contextos de maior variabilidade de valores de *ratings*, também, podemos recorrer à análise da sua relação com as taxas médias de audiências dos suportes para o total da TV. Verificámos que, por vezes, existe uma considerável correlação linear entre estes dois indicadores. Por conseguinte, em determinadas faixas horárias a previsão de valores de *Tvr%*, que será aparentemente mais fácil de realizar, e a posterior aplicação de modelos de Regressão linear ou de Regressão não linear poderá ser a melhor opção. Podemos recorrer, por exemplo, às Árvores de Regressão (*De Ville, 2006; Wei-Yin Loh, 2011*), existindo inúmeras tipologias de algoritmos nesta área.

Apresentam-se algumas sugestões ao nível da viabilidade da previsão das taxas médias de audiências para o total da TV nos seguintes *itens I a IV*.

I. Previsão de $Tvr\_Período\%$ simbólica e por análise das séries temporais

Com base nas análises de segmentação que foram realizadas, avaliou-se a variabilidade bimensal e interanual deste indicador. Frequentemente, na prática, considera-se o valor médio dos seus valores históricos para idênticos períodos horários como o valor da sua previsão. Como foi possível constatar nas análises realizadas, ano a ano, os seus valores médios diferenciam-se no decorrer do tempo, assim como quanto à sua evolução e grau de variabilidade, nem sempre pertencendo aos mesmos agrupamentos horários.

Também, identificámos as faixas horárias que apresentam uma maior ou menor propensão para apresentarem diferentes valores de $Tvr\%$ no decorrer dos anos, estando os valores mais estáveis associados a períodos lazer ou de ausência de uma actividade laboral para a maioria da população activa. Também, existe alguma diferenciação entre a sua variabilidade para o período semanal e para o fim-de-semana. Atendendo à evolução deste indicador verificamos que existem alguns segmentos de telespectadores com diferenciados comportamentos de consumo televisivo, sendo os cenários possíveis os seguintes, desde que considerando os dias laborais:

- Ligar o televisor para assistir apenas ao espaço informativo da noite. Verificamos a maximização dos valores de $Tvr\%$ no decorrer dos espaços informativos e para os quais contribuem fortemente o canal RTP1, existindo um considerável decréscimo neste indicador, logo após a sua conclusão (*Segmento I*).
- Ligar o televisor para assistir ao espaço informativo da noite e aos seus programas sucessores. Verificamos que por efeito de herança, as audiências podem-se manter até sensivelmente às 23H, em função dos factores de disponibilidade (*Segmento II*). Ao fim de semana este período é um pouco mais prolongado.

As selecções programáticas resultam da interacção existente entre a passividade da audiência, as preferências e os vários comportamentos estruturais. O planeamento da grelha e correspondentes espaços comerciais podem influenciar a variabilidade observada na duração da exposição, condicionando a ordem de grandeza dos valores de *ratings*

- Ligar o televisor em vários períodos no decorrer do dia em especial para ver o espaço informativo da tarde e, posteriormente, em maior grau com uma tendência ascendente após as 16.30H (*Segmento III*).
- Ligar o televisor, pontualmente, para ver os programas desportivos ou outros eventos no decorrer da semana, em função dos factores de disponibilidade, e com uma grande variabilidade ao nível da exposição e provavelmente, também, ao nível das selecções (*Segmento IV*).

Ao fim de semana, o primeiro segmento em princípio por seleccionar este *media* como um veículo de informação, também, deverá apresentar um comportamento sensivelmente equivalente. No segmento II, os telespectadores podem manter o mesmo perfil de exposição ou verificarem-se algumas variações ao nível do horizonte temporal que abrange os programas sucessores e que será função das variáveis sociodemográficas ou psicográficas que os caracterizam. Por exemplo, os segmentos mais jovens com um maior interesse em actividades fora do lar devem apresentar uma menor exposição televisiva. Relativamente ao terceiro segmento, também, se poderá manter ou existir uma menor exposição televisiva ao fim-de-semana.

Para os telespectadores que apresentam um comportamento mais aleatório e variável (*segmento IV*) será mais difícil prever a sua exposição televisiva em períodos de lazer. Considerando os primeiros três segmentos será mais fácil conseguir caracterizar e prever os seus comportamentos. Para prever a exposição televisiva e, posteriormente, as selecções programáticas de cada um destes segmentos vão ser necessárias diferentes abordagens e instrumentos analíticos.

O segmento IV será o que contribuirá em maior grau para a variabilidade das taxas médias de audiências para o total da TV, sendo, também, mais difícil prever as suas selecções programáticas. Para os segmentos II e III será relevante ter em atenção os comportamentos de lealdade e os efeitos de herança que influenciam o horizonte temporal em que apresentam o televisor ligado e as suas selecções programáticas. O segmento II, provavelmente, poderá apresentar uma motivação com um maior cariz instrumental relativamente o segmento III. O primeiro segmento deverá apresentar uma maior estabilidade do consumo televisivo.

O painel contribui de forma diferenciada para a evolução deste indicador. As análises não devem ser realizadas a nível individual, mas, sim, considerando o colectivo. Nos capítulos seguintes, é proposta uma sequência de análises que poderá contribuir para os modelos de exposição, considerando os fluxos de audiências, mas que apenas deverá ser privilegiada quando se verificar a inadequabilidade de abordagens mais simples e que requerem um menor investimento temporal e financeiro. As análises de previsão univariadas de *ratings* e, posteriormente, as multivariadas devem ser previamente consideradas.

Ia. Análises de Regressão Simbólicas – Previsão de *Tvr_Período%* incluindo variáveis explicativas e considerando os períodos em princípio mais estáveis

Para prever este indicador, e em princípio apenas nos períodos que apresentam uma menor variabilidade interanual, podemos recorrer à Regressão linear Simbólica considerando como variável dependente *Tvr_Período%* (do ano N) que será da tipologia intervalo e como variáveis dependentes, por exemplo, as seguintes:

➤ Cada faixa horária apresenta uma distribuição de tipologias programáticas, por exemplo. 20% informação, 30% telenovela, 10% desporto, etc. A variação ao nível da composição tipológica no decorrer dos anos poderá influenciar este indicador.

Consequentemente, devemos considerar estas variáveis diagrama que, também, podem representar, de ano para ano, a variação existente nas frequências observadas de cada uma das tipologias programáticas para os diversos canais. As alterações ao nível das tipologias programáticas podem como consequência das preferências dos telespectadores e da influência dos factores estruturais levar a alterações nas taxas médias de audiências para o total da TV. No entanto, estas categorizações impostas pela indústria televisiva, nem sempre, correspondem à segmentação realizada pelo público. Os programas inseridos na mesma designação não têm porque apresentar desempenhos equiparáveis.

➤ A composição sociodemográfica da faixa horária também é relevante, (ex. f1% mulheres, f2% classe social D, ... etc.) em especial considerando os segmentos que apresentam uma maior discrepância na sua evolução face à média observada (variável histograma). Os segmentos que apresentam exposições mais diferenciadas foram identificados nas análises sociodemográficas realizadas aos indicadores de audimetria. Podemos considerar o diferencial de valores entre anos consecutivos, por exemplo, +X% de mulheres em 2009 face a 2008 no decorrer da faixa horária Y, etc.

➤ Valores das taxas médias de audiências para o total da TV para os anos anteriores (variáveis intervalo). Não se deveria seguir a mesma abordagem quando as faixas horárias de, ano para ano, alteram o segmento horário a que pertencem. Quando mantêm a sua classificação podemos recorrer ao seu intervalo de valores para os anos antecedentes. Nas restantes situações outras abordagens devem ser definidas.

➤ Também, é relevante a variação ao nível da composição tipológica e sociodemográfica das faixas horárias adjacentes (períodos horários designados de “Lead_In” e de “Lead_Out”). Verificamos que os canais privados apresentam uma maior variabilidade de valores no decorrer dos anos. Logo, diferentes ponderações têm que ser atribuídas às variações observadas em função dos canais.

➤ A utilidade associada a determinada faixa horária que será função da correspondência existente entre as variáveis sociodemográficas e as tipologias transmitidas. Com base nos índices de afinidade e de atractividade, também, podemos considerar a distribuição de valores para as faixas horárias consideradas. Por exemplo, em 2010, entre as 20H-20.30H, existe um intervalo de valores de atractividade em função dos programas transmitidos para cada um dos segmentos e para o universo que deve ser considerado.

➤ Uma vez caracterizado o perfil sociodemográfico de quem apresenta uma maior ou menor lealdade a determinados, canais e tipologias podemos considerar a percentagem de telespectadores que são leais, ou não, à faixa horária em estudo.

Muitas outras hipóteses podem ser consideradas. Os modelos causais devem incluir as variáveis que podem ser conhecidas no momento de realizar a previsão e devem ser considerados nos contextos de maior estabilidade dos dados, viabilizando a Análise Simbólica a identificação dos intervalos de valores futuros mais prováveis de ocorrer para determinada combinação dos valores das variáveis explicativas consideradas.

Ib. Previsão de valores de $Tvr\_Período\%$ com base em séries temporais ao minuto (períodos mais estáveis)

Também podemos considerar o *clustering* das séries temporais. Por exemplo, considerando os agrupamentos de períodos horários com 30 minutos de duração e identificados anteriormente, devem ser calculadas as correspondentes séries temporais ao minuto ou para unidades temporais um pouco mais abrangentes. Obtemos, assim, uma série temporal para cada dia, $d=1, \dots, m$ em função das diversas faixas horárias. Posteriormente, procedemos ao *clustering* das mesmas, seleccionando a medida de dissimilaridade e o algoritmo de *clustering* mais apropriado. Associamos a cada dia, o segmento a que pertence, por exemplo, 1,2,3,4. Construindo uma janela temporal podemos procurar todas as sequências de séries temporais que apresentem a mesma ordenação de agrupamentos face à série de referência.

A previsão poderá corresponder à média de todas as séries às quais se associam idênticas sequências (Fig. VI.3.2.1). A previsão para o dia d , neste caso, corresponderia à média de x e y .

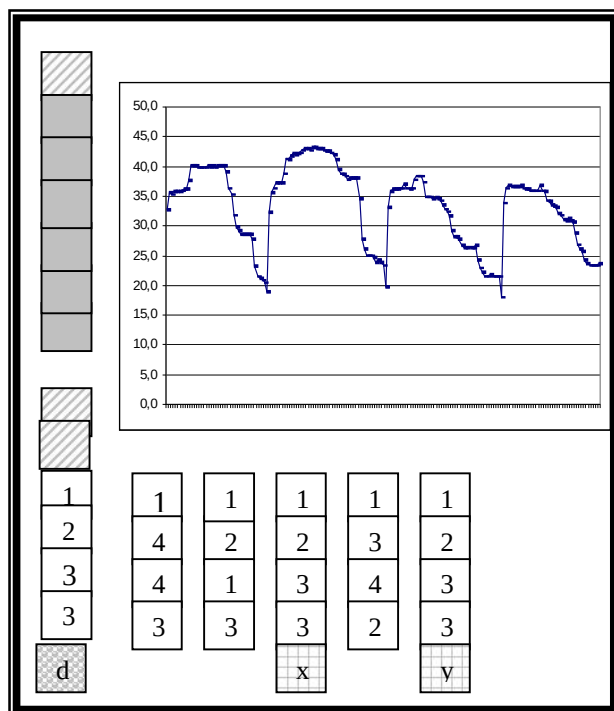


Fig. VI.3.2.1 – A previsão das taxas médias de audiências para o total da TV.

Fonte: Elaborado pela autora.

Nas análises realizadas consta a aplicação de algoritmos de Regressão local linear e local não linear, assim como a segmentação das séries temporais que permitiu alcançar alguns resultados que podem ser considerados satisfatórios.

Verificamos que considerando os dados temporais, as evoluções que se afastam em maior grau da média observada corresponde à transmissão de programas não habituais.

II. A previsão de valores de $Tvr\_Período\%$ para observações atípicas

Para os eventos desportivos será especialmente difícil conseguir prever qual será o acréscimo que se verificará nas taxas médias de audiências para o total da TV. Os *ratings* e os valores de $Tvr\_Período\%$ dos Jogos de futebol são função de inúmeras variáveis entre as quais: faixas horárias, equipas envolvidas, canal, relevância do jogo, etc., entre outros factores qualitativos que podem ser definidos com base nos registos audiovisuais associados aos dados de audimetria. Porém, os modelos causais dificilmente devem conduzir a resultados satisfatórios, atendendo às variáveis não quantificáveis e previsíveis que influenciam a exposição.

Em função das variáveis qualitativas seria necessário calcular quem poderá ligar o televisor que em geral não o faz habitualmente para esse período horário, o que será função de inúmeros factores pessoais que condicionam a disponibilidade de cada um. Para além da disponibilidade, o conhecimento da grelha de programação, também, é relevante, assim como a influência do agregado familiar, entre muitos outros factores. Também, seria necessário a avaliação da sua repercussão nas faixas horárias adjacentes, o que dependerá da disponibilidade, efeitos de herança e da estruturação da grelha de programação. Nestes contextos o que condiciona a viabilidade das análises e os resultados depende em maior grau de conseguir quantificar inúmeros factores de natureza “não previsível” em detrimento das abordagens metodológicas.

Cada jogo apresenta as suas próprias especificidades e características que lhe são particulares e não generalizáveis a outros contextos. Segundo a bibliografia da economia desportiva, dificilmente, podem ser identificadas variáveis que possam ser incluídas em modelos de exposição causais. Logo, é necessário identificar perspectivas alternativas de análise. O mais relevante, nestes contextos, será avaliar a evolução mais provável dos dados face aos valores históricos dos suportes horários em análise para se poder negociar com as estações televisivas e conseguir melhores preços para os suportes publicitários. O *Data Mining* temporal poderá contribuir para a interpretação da evolução dos dados.

Esta questão foi introduzida no *Capítulo VI.3.2.1.* e será abordada num capítulo subsequente, após a descrição das perspectivas de análise mais adequadas aos contextos de uma maior irregularidade dos dados.

III. Previsão com base em dados dos ficheiros de base de audimetria (Em princípio para as faixas horárias onde é superior a variabilidade dos dados)

Verificamos que algumas faixas horárias em dias laborais apresentam uma maior variabilidade de valores no decorrer dos anos.

Para as faixas horárias indicadas como mais variáveis nas análises anuais e para os segmentos que apresentam uma menor consistência e regularidade na exposição televisiva devemos em princípio recorrer aos dados dos ficheiros de base de audimetria (*Quadro VI.3.2.1*).

A variável *Tvr_Periodo%* corresponde ao quociente entre o total de segundos contactados no decorrer do evento ou período horário que se encontra em denominador e o horizonte temporal de referência multiplicado pelo número de telespectadores. Na presente tese, adoptou-se um valor de 30 minutos, isto é:

$$Tvr\_Periodo\% = \frac{Total\ Segundos\ Contactados}{30Minutos \times N^o\ Telespectadores}$$

IIIa. Previsão com base em dados dos ficheiros base de audimetria (Períodos mais estáveis)

Por exemplo, uma estimativa do valor de *Tvr_Periodo%* para as 19H-19.30H de uma Segunda poderia ser o somatório do número de minutos médio que cada telespectador apresenta o televisor ligado no decorrer do intervalo de 30 minutos a dividir pelo produto entre os trinta minutos e o número de telespectadores utilizado no cálculo. Será necessário considerar um elevado horizonte temporal, uma vez que os comportamentos não apresentam ciclos de curta duração. Considerar a evolução média apenas poderá ser viável em contextos de maior estabilidade, uma vez que os comportamentos não são regulares.

Alternativamente, também, podemos utilizar as variáveis simbólicas e calcular um intervalo para o valor das taxas médias de audiências para o total da TV. A exposição ao *media* será quantificada com variáveis intervalo para determinada faixa horária e dia, por exemplo: Telespectador i - [α_i , β_i minutos], ... , Telespectador z - [α_z , β_z minutos].

O total de Segundos Contactados (G) = $[\alpha, \beta]$ minutos] poderia ser o valor resultante da união dos intervalos de valores associados a cada um dos telespectadores: α seria o menor valor dos α_i e β seria o maior dos valores β_i . Utilizando α e β calculamos, também, um intervalo de valores para $Tvr\_Período\%$ (Quadro VI.3.2.1).

IIIb. A previsão de valores de $Tvr\_Prog\%$, isto é, dos valores associados às durações programáticas

Os factores que apresentam repercussão na ordem de grandeza dos valores programáticos são inúmeros, entre os quais os seguintes:

Faixa Horária	$Tvr\_Período$ (%)	Tvr_Prog (%)
1	$\Delta t_{i1}=[\alpha_i, \beta_i]$	$\alpha = \min \alpha_i$ $\beta = \max \beta_i$
2	$\Delta t_{i2}=[\alpha_i, \beta_i]$	
3	$\Delta t_{i3}=[\alpha_i, \beta_i]$	
4	$\Delta t_{i4}=[\alpha_i, \beta_i]$	
5	$\Delta t_{i5}=[\alpha_i, \beta_i]$	

- Alterações horárias na transmissão dos programas. Quanto maior for a diferenciação ao nível dos agrupamentos horários maior será a variabilidade decorrente deste factor;
- Alterações na disponibilidade do painel ou de exposição ao meio televisivo, uma vez que os telespectadores podem estar no lar, mas optar por realizar outras actividades;
- Alterações na duração dos programas;
- Repercussão da programação que gera maiores níveis de exposição.

Quadro VI.3.2.2 – Cálculo de valores de $Tvr\_Prog\%$ a partir dos dados para períodos horários.
Fonte: Elaborado pela autora.

Telespectador I...Z	Dia	Minuto a minuto no decorrer da emissão diária		
Telespectador i		Minuto 1		Minuto M
	Segundas			
	Dia 1	1		0
	Dia N	0		1
....				
Telespectador z				
	Segundas			
	Dia 1	1		0
	Dia N	0		1
				

Variáveis binárias

Telespectador I...Z	Dia	Períodos Horários de 30 minutos		
		19H-19H.30		
Telespectador				
	Segundas			
	Dia 1	N.º minutos que tem o TV ligado		
	...	...		
	Dia N			
	Nº_Min_Médio	N.º Minutos Médio para todas as Segundas		
	Quantil_Médio	1Q – <25%,...4Q->75%		
	Intervalo	Ex. [15, 20 minutos]		
Telespectador z				
	Segundas			
	Dia1	N.º minutos		
	Dia N			
	Nº_Min_Médio	N.º. Minutos Médio para todas as Segundas		
	Quantil_Médio	1Q – <25%,...4Q->75%		
	Intervalo	Ex. [10, 20, minutos]		
Segundas (totalidade telespectadores ou por clusters)				
	Intervalo	[15-25 minutos]		
	Média	N.º de Minutos Médio		
	Q Médio	3Q		
	Variável Diagrama	f1*Q1, f2*Q2, f3*Q3, f4*Q4 sendo os fi as frequências, as % de telespectadores que assistem a determinado quartil		

Quadro VI.3.2.1 – Previsão das taxas médias de audiências para o total da TV.

Fonte: Elaborado pela autora

Para uma grelha de programas sem observações atípicas, uma vez calculado um intervalo de variação para $Tvr\_Período\%$, o valor de $Tvr\_Prog\%$ poderá ser calculado ponderando os valores associados aos suportes temporais que este abrange, como se exemplifica na seguinte *Figura VI.3.2.2* e *Quadro VI.3.2.2* em que $\Delta_{ti}=[\alpha_i, \beta_i]$ representa os diferentes períodos de 30 minutos ou menores que apresentam diferentes graus de variabilidade das taxas médias de audiências para o total da TV, associando-se a cada um deles um intervalo de variabilidade de valores.

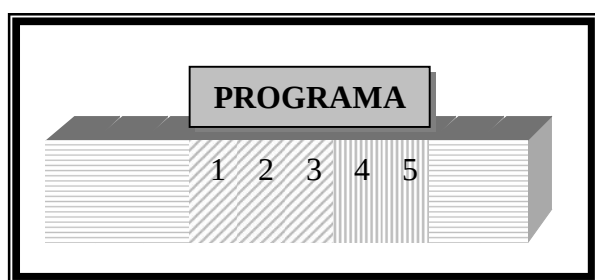


Fig. VI.3.2.2 – Suportes temporais associados a um programa. Fonte: Elaborado pela autora.

A partir dos intervalos de variabilidade de cada período podemos calcular um intervalo de valores para $Tvr\_Prog\%$. Estes valores devem ser calculados com base nas durações programáticas parciais, uma vez que abrangem um menor número de Δ_{ti} .

O limite inferior de $Tvr\_Prog\%$ poderá corresponder ao menor valor de α_i e o limite superior de $Tvr\_Prog\%$ ao β_i mais elevado (*Quadro VI.3.2.2*).

Multiplicando o valor de $Tvr\_Período\%$ pelo número de telespectadores obtemos as taxas médias de audiências do minuto médio, isto é, o número de segundos contactados por parte da totalidade dos telespectadores em cada minuto. $Tvr\_Período\% = \text{Total de segundos contactados} / (\text{período de referência} \times \text{n.º de telespectadores})$ e $Tvr\_Período\_Médio\% = Tvr\_Período\% \times \text{N.º de telespectadores}$

Logo, considerando o exemplo anterior, os valores das taxas médias de audiências para o total da TV podem ser calculados com base na seguinte expressão:

$$Tvr\_Prog\% = \frac{Tvr\_Período\_Médio\%(\Delta_{t1}) * Dur\__{\Delta_{t1}} + \dots + Tvr\_Período\_Médio\%(\Delta_{t5}) * Dur\__{\Delta_{t5}}}{(Dur\__{\Delta_{t1}} + \dots + Dur\__{\Delta_{t5}}) \times \text{Nº de Telespectadores}}$$

Em suma, foi possível verificar a existência de alguma estabilidade na evolução de *Tvr_Periodo%*, se excluirmos programas que podem gerar observações atípicas. Verificamos a existência de alguma independência desta variável face aos conteúdos transmitidos. A sua variabilidade é superior, essencialmente, como consequência das alterações verificadas nos seus horários de transmissão e na duração dos programas e pontualmente pelos conteúdos. Portanto, devemos tentar tirar partido da estabilidade que este indicador, por vezes, apresenta e prever os valores de *ratings* através da modelação da relação entre estas duas variáveis. Independentemente, da perspectiva de previsão adoptada, não será viável alcançar resultados uma elevada precisão. Uma vez previstos os valores das taxas médias de audiências para o total da TV, têm que ser construídos modelos de Regressão que viabilizem o cálculo dos valores de *Ratings*.

Porém, na opinião de alguns especialistas de *Data Mining*, os resultados proporcionados por parte da aplicação de algoritmos mais complexos, em geral, não se deverá repercutir de forma muito significativa na precisão dos resultados alcançados, uma vez que as análises incidem numa área caracterizada por uma forte irregularidade dos dados. No *Apêndice F.6* constam algumas análises que avaliam a relação existente entre estes dois indicadores, do ponto de vista linear e não linear e considerando, quer os valores programáticos, quer os seus valores para as faixas horárias.

V.3.2.3. As análises multivariadas de regressão linear e não linear

Alternativamente, as análises multivariadas, recorrendo aos principais indicadores de audimetria ou a outras variáveis estruturais podem-se basear na aplicação de Árvores e Regras de Regressão, Redes Neurais ou na Regressão linear simbólica, entre outras metodologias. A Regressão linear corresponde a uma abordagem paramétrica que defende a hipótese dos dados seguirem determinado modelo matemático que determina a parte sistemática das observações e a sua componente aleatória.

Os dados resultam de observar variáveis aleatórias de distribuição conhecida, salvo pela presença de alguns parâmetros cujo valor se desconhece e a sua estimação pode ser realizada recorrendo a uma amostra com uma reduzida dimensão. Apesar dos parâmetros apresentarem uma interpretação intuitiva, consiste em uma abordagem de estrutura rígida, e que nem sempre se adequa aos dados. Qualquer função empírica passa próximo dos dados e assume-se o princípio da continuidade quando se interpola ou extrapola. Nas análises de Regressão linear a recta que resume os dados corresponde apenas a uma tendência e apenas nos proporciona um valor em cuja vizinhança o valor real tem mais probabilidade de se encontrar. As análises de Regressão linear não devem ser realizadas para valores que não se encontram no domínio da amostra. Quando as bases de dados apresentam uma elevada dimensionalidade podem ser adoptados os equivalentes algoritmos simbólicos.

Segundo *Zarate (1997)*, Alexis Benguigui, director da ABC, considerado como um dos teóricos com mais prestígio neste âmbito, na década de oitenta, salientou um caminho que deveria ser seguido por ser o mais promissor e que se traduz na utilização de sistemas informáticos de Inteligência Artificial. Lamentavelmente, no final da década de noventa, a sua utilização a nível empresarial ainda era incipiente com apenas a British Broadcasting Channel (BBC) a dar alguns passos nesse sentido por aplicação de Redes Neurais, e que, contudo, apresentavam elevados erros de previsão. Por exemplo, em análises de classificação as taxas de erro apresentam valores de 40%. Porém, a BBC, enfrenta com um painel designado de “instável e dinâmico” e com uma oferta que se altera permanentemente o que não se verifica a nível nacional em todos os contextos e para todos os segmentos. Desde então, em *Weber (2002)*, também, consta a avaliação da precisão dos resultados de previsão recorrendo a Redes Neurais, sendo possível alcançar os 75%.

Porém, as análises não lineares certamente que devem ser privilegiadas. Com base na análise bibliográfica que foi realizada, no âmbito das inúmeras áreas que integram o *Data Mining*, verificamos que já existem contributos e potenciais de aplicação muito significativos e que no futuro ainda mais relevantes podem vir a ser, como, por

exemplo, no âmbito do *Data Mining* Multirrelacional (Dzeroski & Lavrac, 2001; Padhy & Panigrahi, 2012) e da Estatística Relacional (Geetor & Taskar, 2007). Quando a essência do fenómeno que estamos a estudar não se encontra na enumeração de uma série histórica de dados, não é possível realizar previsões através de uma identificação de tendências. A previsão não se deve basear em análises econométricas univariadas e sem ter em conta outras variáveis explicativas externas.

Consequentemente, foram realizadas análises multivariadas lineares e não lineares para avaliar as interdependências existentes entre os diversos indicadores e as variáveis de audimetria, avaliando quais podem ser mais facilmente previsíveis ou serem previamente conhecidas. Verificámos em que medida as abordagens simbólicas e as análises não lineares de *Data Mining* viabilizam melhores resultados ao nível das percentagens de variância explicada, que podem ultrapassar os 80%. Na bibliografia da especialidade as abordagens que são privilegiadas correspondem à avaliação da repercussão das variáveis temporais, assim como do canal e das tipologias na evolução dos indicadores de audimetria.

As análises baseadas na Inteligência Artificial são consideradas como as mais relevantes neste meio. Consistem em dar um corpo de conhecimento ao sistema e este realiza a previsão deduzindo, eventualmente, em cada uma das suas actuações novos conhecimentos e novas Regras, enriquecendo-se permanentemente. O responsável de um canal não poderá sozinho interpretar uma base de dados que apresenta uma elevada dimensionalidade e realizar as suas previsões de acordo com essa interpretação. Os sistemas informáticos são fundamentais, assim como a construção dos modelos de *media*.

Inúmeros autores, defendem que os sistemas existentes não são capazes de ter em atenção a variabilidade dos comportamentos e dos *media* (Makienko, 2012). Logo, são necessários novos caminhos metodológicos que se diferenciem das sondagens de intenção de ver e dos modelos econométricos.

Actualmente, podem ser facilmente ultrapassados recorrendo a metodologias simbólicas, quer multivariadas, quer de análise de séries temporais. As Regras e as Árvores apresentam a vantagem de interpretar os dados de forma mais eficiente que qualquer profissional do meio, sendo de fácil utilização.

Estes métodos enquadram-se nas abordagens não paramétricas, onde podemos incluir o ajuste de modelos paramétricos locais e as técnicas de aprendizagem supervisionada. As Árvores ou as Regras interpretam os dados e produzem um conjunto de conhecimento fundamentado. Correspondem a Regras inteligíveis para os utilizadores e identificam informação que se encontra oculta nas bases de dados. Permitem-nos compreender melhor o comportamento da audiência e realizam uma previsão não linear. As Redes Neurais apresentam um funcionamento opaco, mas as Árvores permitem-nos aceitar ou rejeitar determinados resultados. Verificamos, que numa óptica explicativa, as análises multivariadas podem conduzir a resultados satisfatórios, mas do ponto de vista da previsão abordagens alternativas devem ser avaliadas em determinados contextos.

VI.3.3. A comparação de algumas abordagens analíticas de *Data Mining* e de Estatística multivariada

Na presente tese constam análises de Regressão linear e não linear que visam o desenvolvimento de modelos explicativos de *ratings*, avaliando-se em simultâneo a sua capacidade predictiva. No *apêndice G* apresentam-se algumas considerações a respeito das abordagens metodológicas implementadas. Comparam-se as vantagens e características das metodologias de Árvores de Regressão, as Redes Neurais e a Regressão linear múltipla.

VI.4. A CONSTRUÇÃO DE PLANOS DE *MEDIA* EM CONTEXTOS DE MAIOR VARIABILIDADE DE VALORES DE *RATINGS*

VI.4.1. Considerações a respeito dos comportamentos de consumo televisivo

Quando os suportes são adquiridos no “upfront” ou no “scatter” market é necessário a construção de planos de *media*. Neste contexto, e contrastando com a reduzida bibliografia da especialidade, que aborda essencialmente a análise dos *ratings* ou dos indicadores que são necessários ao seu cálculo, consideramos ser necessário adoptar outra perspectiva de análise.

Quando os horizontes temporais apresentam uma maior variabilidade de valores de *ratings* devem ser adoptadas as análises dos fluxos de audiências considerando os padrões de consumo televisivo por segmentos sociodemográficos e, não, a nível individual e tentando prever a exposição através da análise da estrutura dos comportamentos. Uma maior irregularidade no comportamento agregado, também, dificilmente poderá ser ultrapassada recorrendo aos dados individuais que são classificados como acausais e indeterminados.

O indeterminismo nega que todos os acontecimentos tenham uma causa, logo alguns factos sucedem de forma não “necessária”. Um dos principais argumentos a favor do indeterminismo será a assimetria entre o passado e o futuro, estando este último aberto a influências e, conseqüentemente, não estando determinado, ao contrário, do passado que se encontra fechado. Segundo *Popper (2011)*, um argumento a favor do indeterminismo será a impossibilidade da auto previsão.

Para um sistema ser determinista deverá ser um sistema fechado, que abrange elementos e estados do mesmo tipo ontológico e que apresenta uma sequência temporal e determinadas condições iniciais que são únicas. Portanto, não apresenta como requisito ser previsível, uma vez que os sistemas indeterministas, também o podem ser e, por outro lado, os sistemas deterministas também podem não ser necessariamente previsíveis. No determinismo a ideia de causalidade rege a totalidade do universo, quer dizer, tudo apresenta uma causa, que poderá ser física ou mental, entre outras. A causa poderá ser qualquer coisa e conhecida ou que se possa vir a identificar, ou não, e rege quer as acções humanas, quer os fenómenos naturais.

A maioria da bibliografia identificada defende que os comportamentos individuais são acausais, mas, também, imprevisíveis.

As análises de previsão devem ser viáveis de forma rápida e com base nas análises realizadas verificamos que a audiência é maioritariamente variável na exposição ao *media* e, frequentemente, em menor grau quanto à selecção dos conteúdos, uma vez que grande parte do painel também não tem acesso à TV Cabo. Recorrendo às análises de duplicação das audiências constatamos que, a nível nacional, existe um segmento, que não apresenta uma grande irregularidade na exposição ao meio televisivo.

Nestes contextos de maior regularidade será necessário captar um padrão de exposição e de análise da estrutura dos comportamentos. Também, existem segmentos onde no decorrer do tempo as selecções podem variar, em especial para quem tem acesso à TV Cabo, e a exposição em determinado instante não é considerada, por alguns autores, como sendo relevante para conseguir promover a previsão em instantes posteriores. Na opinião de *Bulck (2006)*, as selecções devem variar e a exposição não é previsível recorrendo à informação que os telespectadores manifestam em momentos anteriores.

Os sistemas de previsão baseados na análise de séries históricas não devem produzir bons resultados nos contextos de maior variabilidade de *ratings* (*Danaher et al., 2011*), mas poderá existir no conjunto dos programas um padrão de ver TV. Para detectar essa possível ordem aparentemente oculta nos comportamentos da audiência será necessário recorrer essencialmente ao *Data Mining* não Proposicional. As selecções podem não coincidir, de semana para semana, para alguns telespectadores, mas poderá existir um padrão de exposição ao meio televisivo.

Na bibliografia da especialidade, constam essencialmente três perspectivas que explicam a exposição televisiva e que defendem a repercussão de diferentes factores:

(1) A selecção de programas visa satisfazer “necessidades”, e, logo, a selecção é intencional e reflecte uma selectividade e utilidade procurada pelos telespectadores;

- (2) Procura-se ver TV, ou seja, prevalece a “experiência televisiva” como forma de preenchimento de um tempo vazio e eventualmente de lazer. Verifica-se uma ausência de selecção programática intencional em que qualquer opção será, igualmente, interessante ou a selecção da “melhor alternativa” disponível nesse momento;
- (3) As teorias estruturais que defendem a forte influência dos factores de planeamento e dos comportamentos que dela resultam no processo de selecção, entre outras variáveis e que as escolhas e as necessidades individuais, também, se tornam irrelevantes neste contexto. A estruturação da grelha produz determinados comportamentos e os restantes factores estruturais determinam os conteúdos a que os telespectadores se encontram expostos.

Consequentemente, podemos identificar diversas situações relativas à exposição televisiva:

- (1) Exposição intencional associada a uma escolha consciente;
- (2) Exposição não intencional, logo, não se importam de ver determinados programas e os factores estruturais colocam os telespectadores em contextos que o viabilizam;
- (3) Evitar de forma intencional determinado conteúdos;
- (4) Evitar de forma não intencional os conteúdos que gostariam de ver, mas que as variáveis estruturais acabam por inviabilizar.

De acordo com a *Figura V.3.2.1* propomos que estes quatro contextos de exposição podem ser ainda desagregados nos seguintes:

- (1) A duplicação de audiências de “lealdade programática” intencional e que resulta de uma procura activa e interesse nos conteúdos, por exemplo, de uma mini-série.
- (2) A duplicação de audiências de “lealdade programática” não intencional e que decorre do hábito e disponibilidade para ver TV em determinadas faixas horárias. Segundo a bibliografia da especialidade, existem dúvidas sobre se esta tipologia de comportamentos corresponde maioritariamente à existência de actividade ou de passividade nas selecções.

Para um determinado segmento, as selecções podem corresponder a uma preferência por determinado conteúdo ou tipologia e ser intencional, mas para os restantes poderá prevalecer a experiência televisiva e a duplicação das audiências decorrer do conjunto de opções televisivas disponíveis nesse momento e da disponibilidade e vontade para ver TV;

(3) Selecção intencional de um programa, como, por exemplo, de eventos desportivos, galas, etc. Por exemplo, a exposição aos espaços desportivos e aos jogos de futebol, com uma repercussão muito variável ao nível da programação adjacente e concorrente. Por vezes, podemos observar a existência de efeitos de herança num mesmo canal ou de fluxos de audiências para os canais concorrentes, mas, também, podem existir alterações na evolução dos indicadores apenas no decorrer desse programa;

(4) Selecção não intencional por efeito de herança, isto é, que pode resultar da passividade e inércia do público. Neste contexto os telespectadores assistem ao programa (ou programas) que sucede ou antecede uma escolha que foi realizada de forma consciente, uma vez que ainda se tem alguma disponibilidade e pretende-se ver TV, estando o suporte eventualmente incluído numa tipologia em que incide maioritariamente o seu repertório programático.

(5) Selecção não intencional que decorre da sua lealdade a um canal, escolhendo-se os sucessivos conteúdos que são transmitidos nesse canal, porque se pretende ver TV e em maior grau um canal em concreto, uma vez que existe uma maior identificação com a sua imagem e características dos conteúdos transmitidos;

(6) Selecção não intencional que decorre da lealdade tipológica, ou seja, ver um programa que está a ser transmitido em um dos canais que consta do seu repertório uma vez que se tem disponibilidade e se identificam com determinada tipologia de conteúdos;

(7) Evitar um programa de forma intencional, uma vez que apresenta um conteúdo ou insere-se numa tipologia ou canal que desagrada o telespectador;

(8) Evitar um programa de forma não intencional, que pode derivar de inúmeros contextos que não são previsíveis. Poderá ser por influência do grupo em que está inserido, por desconhecimento da programação, por ser transmitido num canal que não consta do seu repertório mais habitual, por falta de disponibilidade, entre muitos outros factores.

Atendendo a todos os contextos de exposição que podem surgir, verificamos que as selecções, ao contrário, do que defende a abordagem das gratificações e usos do *media*, não decorrem obrigatoriamente de preferências e interesses. Os telespectadores, também, podem-se diferenciar ao nível de decisão, segundo o qual ligam o televisor e seleccionam os programas. Quer dizer, independentemente, dos factores que condicionam e caracterizam em maior grau o consumo televisivo podemos diferenciar os seguintes níveis de decisão no acesso ao *media*: (1) selecção de canal, (2) selecção de tipologia, (3) selecção de programa, (4) selecção de parte de programa.

A decisão poderá ser ligar a TV para ver determinado canal e ver a programação existente nesse período; ver TV e escolher a melhor alternativa disponível; seleccionar os programas enquadrados em determinadas tipologias e seleccionar alguns programas em concreto. Um mesmo indivíduo poderá em função dos contextos ser mais ou menos intencional na exposição e mais ou menos regular.

Nas situações de intencionalidade, uma vez que as decisões são tomadas em circunstâncias futuras não conhecidas pelo próprio indivíduo, *a priori* inclusive os métodos de previsão mais qualitativos não devem produzir bons resultados. Inclusive quando a exposição é intencional a “vontade” não é previsível antecipadamente, assim como a disponibilidade e, logo, as sondagens de intenção de ver não produzem bons resultados.

Os contextos de não intencionalidade, uma vez que dependem do preenchimento de um tempo vazio ou de lazer, ainda devem ser mais dificilmente previsíveis para alguns segmentos de telespectadores que apresentam ausência de rotinas diárias.

Certamente que para um vasto segmento de telespectadores, as selecções e a duração da exposição variam significativamente no decorrer do tempo. Consequentemente, será mais fácil o desenvolvimento de abordagens metodológicas alternativas que viabilizem o planeamento de *media* do que o cálculo de alguns dos indicadores para suportes em concreto e que são fundamentais para o sector televisivo. Neste âmbito em primeiro lugar será necessário diferenciar os diferentes contextos existentes, por exemplo, em função da existência ou ausência de hábitos de consumo televisivo e que obviamente condiciona o grau de regularidade da exposição aos *media*. No seguinte capítulo constam algumas considerações a respeito do conceito de hábito.

VI.4.2. O conceito de hábito e a segmentação das audiências

O conceito de “hábito” corresponde a uma temática que suscita alguma controvérsia uma vez que em geral é identificado como uma indicação da passividade da audiência. Contudo, também tem sido operacionalizado como uma das possíveis gratificações do acesso ao *media*. Também poderá não ser uma mera tipologia de “usos do *media*”, mas parte de um processo psicológico que origina e mantém todos os restantes comportamentos de exposição ao *media*.

O conceito de hábito foi associado a ambas as principais teorias que explicam a exposição ao *media*, mas a investigação mais recente relaciona-o em maior grau com as teorias estruturais de análise da exposição televisiva, que defendem uma menor intencionalidade subjacente às decisões de acesso ao *media* e de selecção programática. Quem defende uma concepção mais passiva da audiência tem associado o hábito às teorias que veem o *media* como uma mera actividade lúdica que permite preencher os tempos livres ou vazios. Logo, a disponibilidade e, não, a selectividade desempenha um maior poder explicativo da exposição.

Verificámos a existência de uma forte repercussão da variável que representa a exposição global ao *media*, quer na evolução de alguns indicadores de audimetria, quer considerando a duplicação das audiências. O acesso ao *media* poderá ser passivo e a selecção mais activa. Os factores estruturais, como, por exemplo, a lealdade, os efeitos de herança e a disponibilidade, podem ser variáveis com capacidade de previsão e com uma maior relevância do que a tipologia programática.

Para a Psicologia Cognitiva, o hábito tem sido caracterizado com uma forma de minimizar o esforço cognitivo no processamento da informação. Alguns estudos sugerem que os comportamentos diários apenas derivam parcialmente de intenções e podem ser controlados em parte pelo hábito. O hábito tende a ser operacionalizado como o número de vezes que um acto foi praticado no passado.

Consequentemente, independentemente dos factores que condicionam o desenvolvimento dos hábitos de consumo televisivo, ou quaisquer outros, a sua existência ou ausência condiciona as abordagens de previsão a adoptar. Segundo as análises realizadas às taxas médias de audiências para o total da TV, facilmente identificamos pelo menos cinco segmentos de telespectadores com padrões de consumo televisivo diferenciados no decorrer do tempo. Estes segmentos podem apresentar diferentes características sociodemográficas e psicográficas, contribuindo de forma diferenciada para a variabilidade das taxas médias de audiências, em função dos horizontes temporais em análise. Os segmentos divergem quanto à eventual existência de hábitos de consumo televisivo (*Fig. VI.4.2.1*).

Em função dos contextos e dos segmentos, diferem as abordagens que podem viabilizar os melhores resultados do ponto de vista da construção dos modelos de exposição. A segmentação dos registos que caracterizam a exposição individual, também, poderá condicionar as análises de previsão para as faixas horárias que apresentam uma maior variabilidade dos indicadores de audimetria.

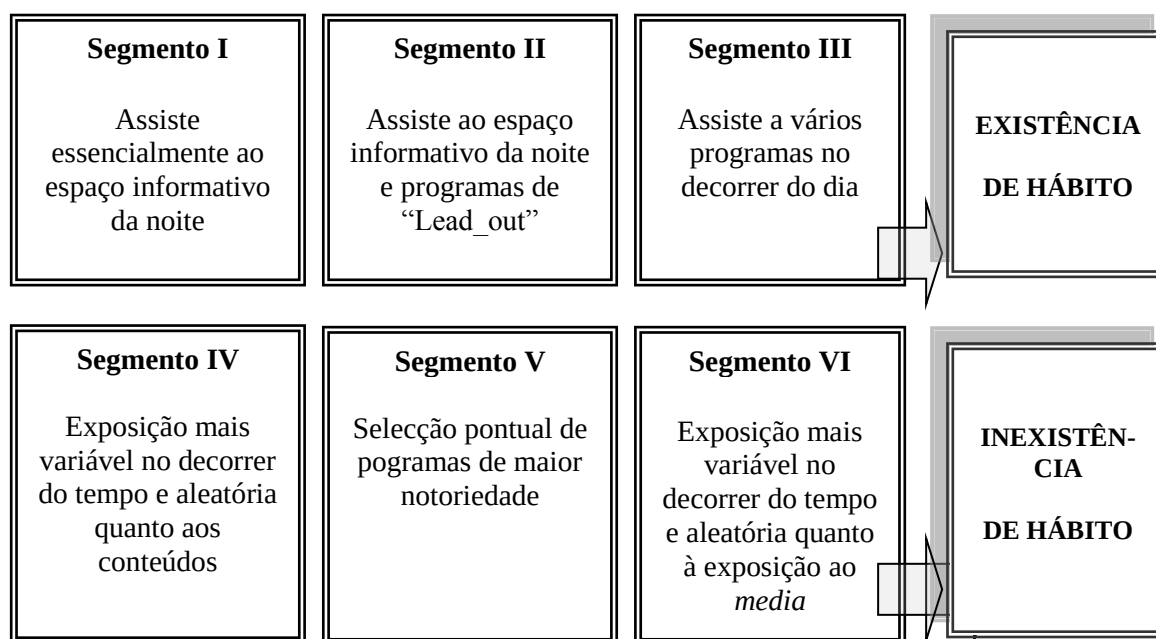


Fig. VI.4.2.1 – Segmentos de exposição televisiva do painel de audimetria.
Fonte: Elaborado pela autora.

As análises de segmentação das bases de dados de audimetria levantam consideráveis dificuldades. De acordo como *Kim (2002)*, tem sido difícil a adaptação das técnicas de segmentação de mercados para a identificação de agrupamentos de telespectadores em função dos seus perfis de consumo televisivo.

Viewing decisions have a direct impact on the *media* planning of broadcasters and consequently they search for segments with similar viewing preferences. If there are segments that make their decisions based on different preferences or tastes, it should be important for *media* planners to know the composition of each segment and its' particular decision process. Until now however there have been several difficulties in adapting market segmentation techniques to find market segments easily in the broadcasting market. (p.51)

A segmentação das audiências televisivas apresenta uma grande relevância para a construção dos modelos de exposição, assim como para a definição do perfil de audiências de canais, tipologias programáticas e programas e para a avaliação do grau de envolvimento dos telespectadores com a grelha de programas. A análise dos comportamentos de lealdade, também, poderá constituir um instrumento de previsão. Por vezes, as decisões de exposição ao *media* podem ser realizadas ao nível dos canais, tipologias ou programas em concreto.

Não existe consenso na bibliografia da especialidade relativamente à heterogeneidade subjacente à exposição televisiva e em que medida as variáveis sociodemográficas na realidade discriminam os padrões de consumo televisivo. Por exemplo, *Cádima (1996)* reconhece a possível existência de cinco categorias de público:

Os que não veem televisão, o público da manhã, o público da hora do almoço, o auditório de fim de tarde e os *heavy viewers* que apenas veem à noite. Estas categorias conferem uma contradição à televisão generalista, ou seja, mesmo nos programas destinados a um público minoritário é a maioria que os vê. (p.58)

Segundo *Fernandes (2001)*, nas entrevistas realizadas nas estações de televisão, verificou que existe algum consenso no que diz respeito ao conceito de público de televisão: não existe um público, mas vários, nomeadamente, segundo a faixa etária e a classe social. Por outro lado, para *Souchon (1990)*, apenas se destacam dois grupos: um público de longa duração e um público que vê pouca televisão. O autor salienta que público de longa duração vê televisão para se informar, distrair, saber notícias do que se passa no mundo, pois não dispõe de outros meios de lazer, cultura e informação. O público que vê pouca televisão atribui à TV uma função acessória, como, por exemplo, o caso dos jovens que saem muito, principalmente nas cidades onde existem alternativas culturais e desportivas.

Na opinião de *Cádima (1996)*, verifica-se uma indiferenciação no plano da segmentação dos públicos, ou seja, não existem públicos específicos, mas, sim, um grande público. “Embora se considere (e as pessoas entrevistadas na televisão referiram esta questão da segmentação do público) que existem vários públicos, Cádima alerta que as teorias recentes têm desmentido “esta verdade quase consagrada.” (*Fernandes, 2001, p.124*).

A mesma conclusão, também, é referida em *Huertas (2002)*, que apresenta as variáveis sociodemográficas como insuficientes para diferenciar os padrões de consumo, apresentando os diversos segmentos sociodemográficos comportamentos na realidade consideravelmente semelhantes.

Salienta que, no geral, as identidades grupais são cada vez mais apenas provisionais, e que as suas fronteiras são difusas, ou seja, as pessoas podem pertencer simultaneamente a vários grupos e, frequentemente, assistimos à sua passagem entre eles. Os instrumentos metodológicos utilizados para a construção dos modelos de exposição devem ter em consideração esta tendência, ou seja, a instabilidade dos segmentos no decorrer do tempo.

Com base na bibliografia da especialidade, essencialmente, de cariz mais qualitativo verificamos que não existe qualquer consenso relativamente ao número e características dos segmentos que possam existir, assim como quanto à relevância das variáveis sociodemográficas para a sua caracterização.

As sociedades cosmopolitas e sofisticadas podem gerar diversos estratos que frequentemente retrocedem e procuram uma identidade simplificadora. Uma maior oferta televisiva, também, nem sempre incute uma grande variabilidade ao nível dos conteúdos seleccionados e reduz o número de segmentos com comportamentos muito diferenciados. O perfil das audiências, por sua vez, também é determinante para o planeamento de *media*.

Em *Kim (2002)*, no âmbito da investigação quantitativa consta uma sugestão metodológica para a identificação de segmentos no mercado televisivo Coreano, recorrendo a índices de correlação e ao algoritmo K Média, que apenas poderá ser implementado se os dados dos ficheiros de base de audimetria estiverem disponíveis. Inevitavelmente terá que existir uma forte divergência de opiniões, em função das abordagens analíticas e metodologias de investigação adoptadas. Consequentemente, uma vez mais, a Estatística condiciona a segunda área de que depende fortemente o planeamento de *media*, ou seja, a segmentação das audiências televisivas.

A nível empresarial, frequentemente, as previsões são realizadas por público objectivo, os quais se podem diferenciar nas abordagens mais adequadas, contudo na prática nem sempre se tem este aspecto em consideração.

Também são necessárias abordagens metodológicas alternativas considerando os indicadores agregados, uma vez que em geral os utilizadores destes dados não têm acesso aos dados dos ficheiros de base de audimetria (*Danaher et al. 2011*) e que corresponde aos dados que são utilizados em *Kim (2002)*.

Quando se constrói um plano de *media* é necessário seleccionar os suportes que apresentam uma maior adesão por parte do público-alvo, segmentar as audiências e compreender o seu processo de decisão. A introdução de um novo programa torna necessário avaliar quais são as faixas horárias em que o perfil do seu público poderá apresentar uma maior predisposição para o incluir no seu repertório programático. As faixas horárias em que a programação concorrente apresente por parte do seu público fortes comportamentos de lealdade poderá condicionar fortemente o seu desempenho.

Nesse sentido, com base no *Data Mining* e na Análise Simbólica apresentam-se algumas alternativas metodológicas que permitem a definição do perfil de audiências dos canais, programas e tipologias programáticas, considerando os indicadores que caracterizam a lealdade programática (*Apêndice F.5*). O contexto de “estruturação das opções televisivas” referido em *Kim (2002)* permite explorar as relações existentes entre os suportes e as variáveis que caracterizam o painel.

Na área da segmentação devem ser avaliadas as melhores alternativas, quer a nível agregado, quer considerando os dados dos ficheiros de base de audimetria. Em *Kim (2002)*, descreve-se uma abordagem metodológica de segmentação dos dados, salientando o autor que as selecções dos telespectadores são realizadas com base em “preferências”. As teorias estruturais da exposição televisiva têm produzido melhores resultados no âmbito do desenvolvimento de modelos explicativos e de previsão de audiências televisivas, do que as teorias que defendem as selecções como uma função das preferências programáticas ou tipológicas.

Atendendo às análises realizadas, a evolução dos *ratings* face a uma grelha de programação já estabelecida encontra-se, essencialmente, condicionada pela evolução de *Tvr_Período%* que apresenta alguma regularidade e independência face ao conteúdo transmitido e por parte do grau de duplicação de audiências num mesmo canal, em detrimento dos fluxos entre canais. A duplicação de audiências encontra-se por sua vez dependente das decisões de ligar e desligar o televisor, que será função em maior grau dos factores de disponibilidade, assim como também do indicador sociodemográfico que representa a fidelidade programática. De acordo com *Copper e Tang (2009)*, as audiências são activas em ambientes estruturais, logo as selecções encontram-se influenciadas pelas estruturas subjacentes e pelos comportamentos que resultam da forma como as grelhas de programação são estruturadas. Assim, e coincidindo com *Souchon (1990)*, verificamos que devem existir dois grupos: um público de maior duração e um público que em geral assiste a pouca televisão, com uma exposição mais variável. Inclusive nos períodos de maiores audiências, a duplicação de audiências de lealdade programática não atinge o valor de 50%.

Contudo, e de acordo com Cádima, verificamos que neste segmento caracterizado por uma maior exposição ao *media*, existem diversos subsegmentos. Identificamos um segmento de “Heavy viewers” que apenas assistem à TV no período da noite, mas os restantes segmentos que o autor descreve, não parecem se adaptar à realidade nacional, pelo menos para o período em análise e com base nos resultados das análises de duplicação de audiências que foram realizadas. De acordo com a *Figura VI.4.2.1* foi possível, sim, identificar os seguintes segmentos de telespectadores:

- (1) Quem assiste apenas aos espaços informativos da noite;
- (2) Quem assiste aos espaços informativos da noite e a alguns dos seus programas sucessores e, eventualmente, também aos antecessores, mas com uma menor duração de exposição. A duração da exposição será função dos factores de disponibilidade e vontade, permanecendo alguns telespectadores sempre no mesmo canal e outra fracção transitando entre canais, em especial entre os canais privados generalistas, para ver os programas que considera como sendo “mais apelativos” atendendo à oferta disponível, ou sendo a exposição decorrente da influencia de determinados factores estruturais;

- (3) Um segmento que em função dos factores de disponibilidade assiste à TV no decorrer de todo o dia em diferentes sessões, assistindo nomeadamente ao espaço informativo da tarde e a partir das 16.30H apresentando uma tendência crescente de duplicação das audiências;
- (4) Também existe um segmento que apresenta uma maior exposição no período matinal.

Se nestes segmentos existe uma determinada componente de hábito, também, podemos identificar outros em que isto não se verifica. *Souchon (1990)* refere que o público que vê pouca televisão atribui à TV uma função acessória. Os eventos desportivos apresentam uma forte diferenciação quanto à sua capacidade para cativar uma parte deste público. Para este segmento será especialmente difícil conseguir prever quais são as suas selecções e a exposição ao *media*, inculindo uma maior variabilidade aos indicadores de audimetria. Alguns telespectadores apresentam uma exposição mais variável aos conteúdos, e, por vezes, inclusive ao nível das faixas horárias, mas existem outros segmentos que não apresentam o hábito de ver em determinadas faixas horárias, ligando o televisor para ver, essencialmente, determinados eventos, como, por exemplo, os programas desportivos.

VI.4.3. A selecção de suportes

Na literatura da especialidade foram identificados alguns escassos trabalhos que abordam a análise de alguns dos aspectos enumerados anteriormente, recorrendo em geral a metodologias de análise estatística mais tradicionais. Nomeadamente, em *Kim (2002)*, consta um processo de segmentação das audiências televisivas e de selecção de suportes. Inicialmente, o autor procede à recolha dos dados dos ficheiros de base de audimetria. Segue-se o cálculo de um índice de correlação que avalia a probabilidade de determinado par de programas vir a ser seleccionado por um telespectador e a aplicação da análise de *clusters* para identificar sub-mercados de programas. Por último, verifica-se a aplicação do algoritmo K Média para identificar os segmentos de telespectadores, utilizando como semente os segmentos de programas identificados anteriormente e conclui-se com a análise do perfil sociodemográfico de cada segmento identificado.

Como alternativas metodológicas que viabilizam a identificação e caracterização de segmentos de telespectadores e que podem ser utilizadas em determinadas etapas de construção dos planos de *media* (em especial quando as análises incidem nos segmentos que não apresentam “hábitos” de consumo televisivo), nomeadamente para uma pré-selecção dos suportes, referem-se as seguintes duas opções:

Ia. Segmentação e selecção de suportes recorrendo aos dados dos ficheiros base de audimetria

ETAPA I – *Clustering* de registos de exposição televisiva, identificando os registos mais semelhantes e representativos da exposição de cada telespectador (*Fig VI.4.3.2*) e por segmento (*Fig. VI.4.3.1*) identificado, por aplicação de algoritmos de *clustering* de bases de dados transaccionais (*Ketterlin, 1997; Smyth, 1999*), como, por exemplo, os algoritmos de densidade.

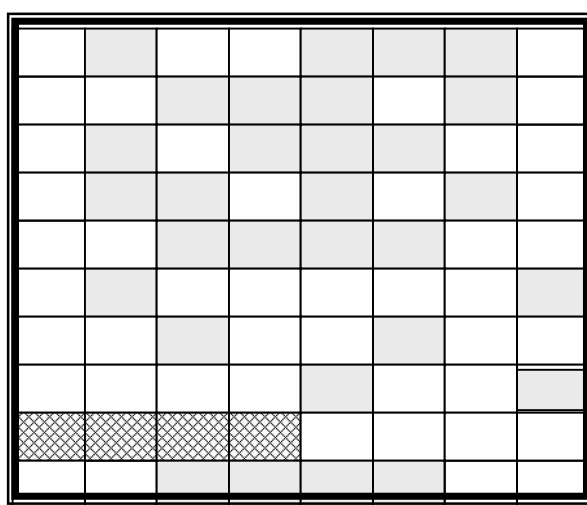


Fig. VI.4.3.1 – Identificação das sequências de exposição mais características por telespectador
Fonte: Elaborada pela autora.

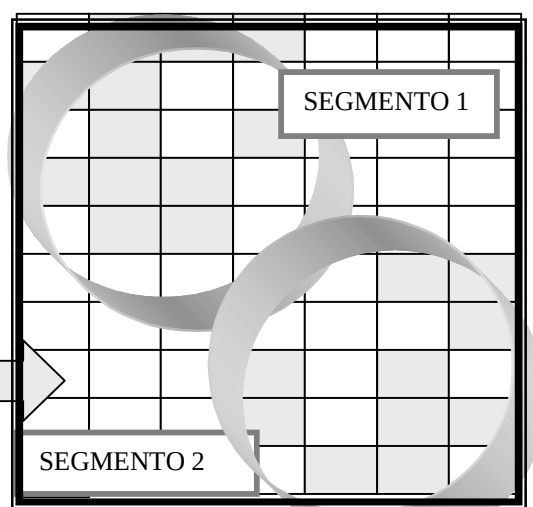


Fig. VI.4.3.2 – Identificação dos segmentos de telespectadores por *clustering* de sequencias de exposição. Fonte: Elaborado pela autora.

Implicitamente, também, podem ser identificadas algumas associações entre os suportes, nomeadamente os que realmente se complementam e não resultam de uma exposição meramente pontual e pouco representativa do consumo televisivo.

Pretende-se a aquisição dos suportes que estejam associados a um eventual padrão de exposição televisiva e nos quais se dupliquem as audiências, sendo o valor mais adequado da sua intensidade uma função das características do produto, serviço, mercado da campanha publicitária e do seu público-alvo.

ETAPA II – Derivar Regras de Associação (*Zaki, 2001*) ou Sequenciais (*Xin, Han, Cheng & Yan, 2007*), verificando para cada segmento previamente identificado quais são as associações mais características entre os suportes em análise, sejam estes programas ou períodos horários. Para além de identificar quais são os suportes a que os telespectadores assistem com uma maior frequência nas diversas sessões televisivas, também, podemos identificar quais são os programas que implicam a ausência de selecção de outras tipologias programáticas. Posteriormente, a caracterização de cada segmento recorrendo a variáveis sociodemográficas ou/e psicográficas, por exemplo, por aplicação de metodologias de Análise Simbólica.

Ib. Segmentação das audiências e a estruturação dos suportes recorrendo aos dados agregados, ou seja, aos indicadores de audimetria ou aos dados dos fluxos de audiências entre suportes para cada um dos segmentos.

Podem ser seleccionados para determinados macrosegmentos, os programas que apresentam índices de atractividade ou adesão superiores a um valor definido pelo utilizador e, assim, identificamos os programas que apresentam maiores associações no decorrer de um dia. Os macrosegmentos podem ser identificados com as metodologias simbólicas de *clustering* e de dissemelhança entre objectos simbólicos. Nestes contextos as abordagens metodológicas descritas no *Apêndice F.5*, também, são válidas.

Os eventos vão corresponder a segmentos e aos programas que apresentam valores de afinidade superiores à unidade ou, por exemplo, no 4º quantil para determinado dia, ou seja, a: Segmento_Id, Data, Horizonte Temporal_Y, Canal_X, {Prog_X, Prog_Y, etc. Que representa o conjunto de programas para os quais o segmento identificado apresenta valores de adesão ou afinidade superiores à unidade}.

Também, podemos utilizar variáveis relativas aos canais, como, por exemplo, Canal_X, Data, Horizonte Temporal, {Tempo_Total_Despendido no Prog_X ou na Tipologia_Programática_Y; Ratings para Prog_X ou Tipologia_Y}. Uma vez construída a base de dados, recorrendo ao algoritmo Apriori do *software* SAS ou a qualquer outro que lhe seja equivalente derivamos Regras sequenciais na qual será relevante a ordem dos programas.

Alternativamente, as associações tipológicas e programáticas têm que ser realizadas através da construção de mapas perceptuais recorrendo a metodologias clássicas ou simbólicas. No *Apêndice F.5* constam algumas análises a este nível, demonstrando-se como algumas metodologias de estatística multivariada e de Análise Simbólica permitem identificar o perfil das audiências de programas, nomeadamente a análise de *clusters*, Discriminante linear e Regressão linear que incidem nos resultados das análises MDS.

VI.4.4. Os modelos de exposição em contextos de existência de hábitos de consumo televisivo

Nos contextos em que os telespectadores apresentam uma maior e mais regular exposição ao *media* podemos identificar os suportes a contratar através da sequência metodológica que consta da seguinte *Figura VI.4.4.1*.

VI.4.4.1. Os modelos de exposição para uma sequência de suportes quando existe o “hábito” de ver TV

Para o primeiro segmento identificado na *Figura VI.4.2.1*, existindo o hábito de ver determinado programa e faixa horária, ainda que, por vezes, possa existir ausência de exposição, as abordagens a utilizar serão função do grau de regularidade de exposição. Se for muito regular poderá ser assumida a exposição histórica para o curto prazo. Se existir ausência de exposição com alguma frequência, não será suficiente verificar, com base no historial de selecções, a sua frequência de selecção e a duração histórica mais representativa.

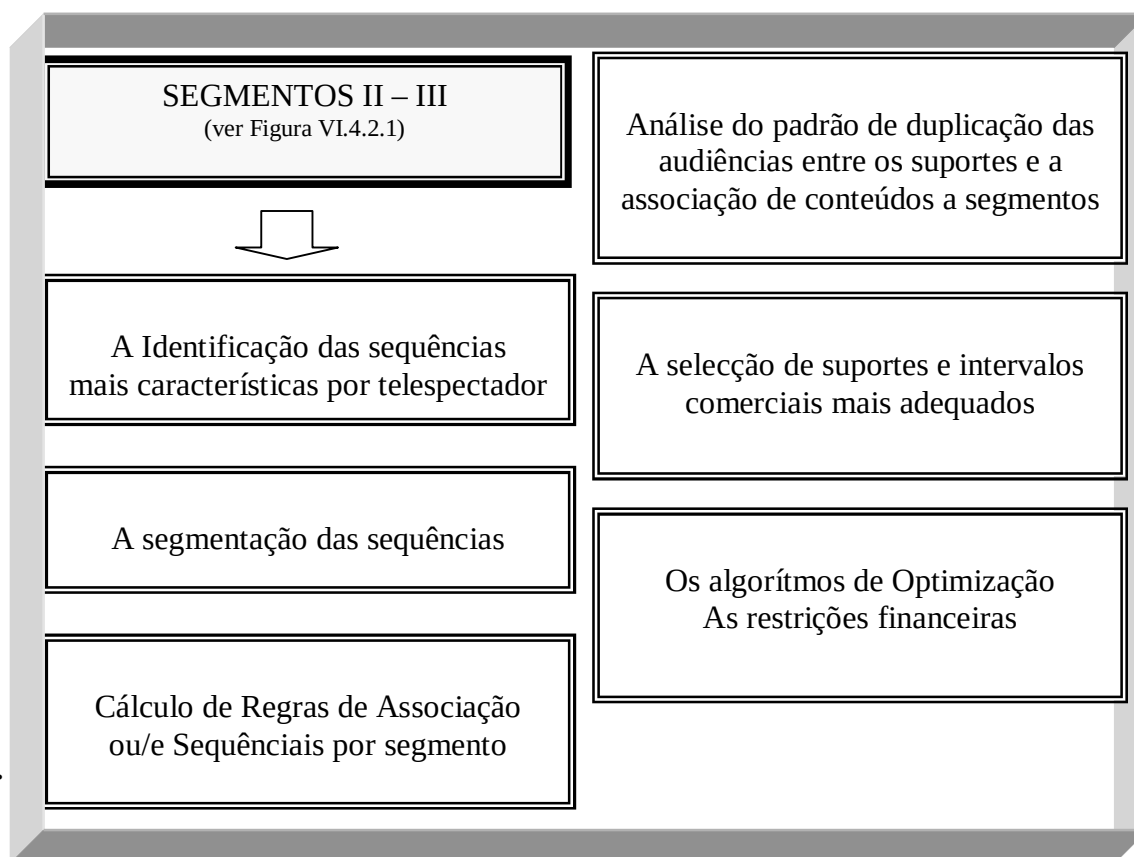


Fig. VI.4.4.1 – Sequência de análises para a identificação de suportes publicitários.
Fonte: Elaborado pela autora.

Ainda que recorrendo a um longo horizonte temporal, essas frequências podem ser variáveis no decorrer do tempo e ser função do horizonte temporal em análise. Em alternativa poderá ser adoptada a análise de médias móveis ponderadas. Para qualquer um dos segmentos que apresenta uma maior regularidade de exposição, esta poderá, talvez, ser prevista por faixas horárias e recorrendo às abordagens metodológicas que se caracterizam nas seguintes páginas se as análises incidirem numa sequência temporal e/ou programática e considerando os dados dos ficheiros de base de audimetria ou os indicadores agregados.

A *Figura VI.4.4.1.1* representa algumas das abordagens que podem ser utilizadas para a construção dos modelos de exposição, quando as análises incidem nos dados do segmento que essencialmente assiste ao espaço informativo, sendo, também, adequada a

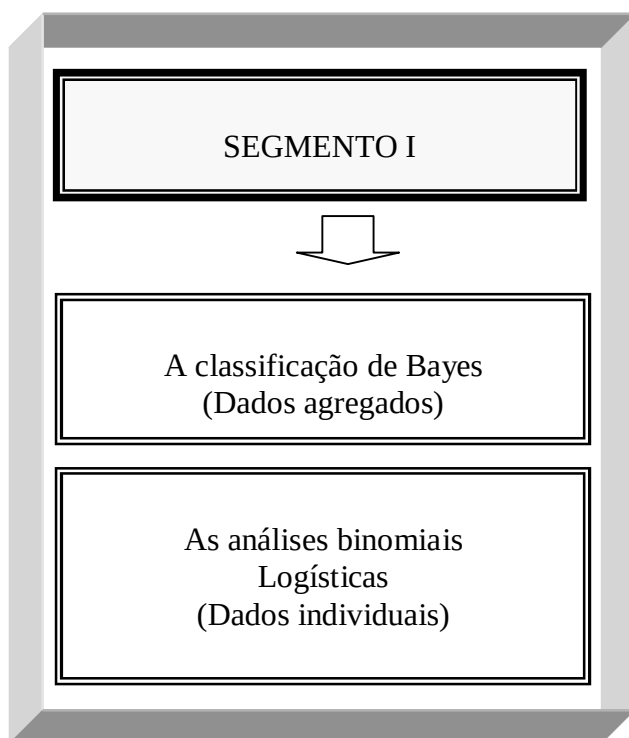


Fig. VI.4.4.1.1 – Sequência de análises para a identificação de suportes.

Fonte: Elaborado pela autora.

outras tipologias de programas que apresentem semelhante comportamento por parte do seu público.

Nestes contextos poderá ser considerada a Classificação de Bayes e as análises Binomiais Logísticas.

I. A Construção de modelos de exposição recorrendo à classificação de bayes

Quando é necessário considerar a influência de algumas variáveis estruturais podemos recorrer à classificação de Bayes (*Becker, Kohavi & Sommerfield, 2001*). A título exemplificativo, no *Quadro VI.4.4.1.1* consta um historial hipotético de selecções programáticas de determinado telespectador.

Considera-se um período de nove dias e sempre o mesmo horizonte temporal. Para definir a variável que representa a selecção do canal podemos recorrer aos dados dos ficheiros base de audimetria, mas, também, aos índices sociodemográficos de adesão ou de atractividade. Tendo como objectivo o cálculo da probabilidade de determinado telespectador ou conjunto de telespectadores seleccionar determinado canal ou programa em determinado dia e para um período horário equivalente podemos recorrer a algoritmos de classificação de Bayes, desde que discretizando as variáveis numéricas ou assumindo uma distribuição Normal.

Dia	Canal	Tipologia SIC	Tipologia TVI	Tipologia RTP1	Valores Tvr%	Synldin_1	SyLdout_2
1	SIC	Telenovela	Noticias	Reportagem	20	v	F
2	TVI	Noticias	Telenovela	Concurso	19	v	F
3	SIC	Telenovela	Telenovela	Concurso	15	v	V
4	SIC	Telenovela	Telenovela	Concurso	12	F	V
5	SIC	Telenovela	Telenovela	Concurso	17	v	V
6	SIC	Telenovela	Telenovela	Debate	21	v	V
7	SIC	Telenovela	Telenovela	Noticias	13	F	V
8	SIC	Telenovela	Filme	Desporto	10	F	F
9	SIC	Telenovela	Filme	Desporto	16	F	F

Quadro VI.4.4.1.1 - Selecções programáticas e a avaliação das influência das variáveis estruturais.
Fonte: Elaborado pela autora.

Um classificador corresponde a uma função $F: \Omega_x \rightarrow \Omega_c$ que atribui uma classe C , que neste caso será o canal, a determinada observação caracterizada por um conjunto de variáveis $X = \{X_1, \dots, X_n\} \in \Omega_x$. Pretende-se calcular a probabilidade $P(C_j|x)$, ou seja, de x pertencer a C_j . A classe a atribuir corresponde à que maximizar as probabilidades *a posteriori*, representadas por $C^* = h_c(x) = \arg \max_{j=1, \dots, n} P(C_j|x)$ sendo $P(C_j|x) = \frac{P(C) \times P(X|C)}{P(X)}$

Para o algoritmo *Naive Bayes* (Langley, Iba & Thompson, 1992), a regra de classificação representa-se pela seguinte expressão:

$$C^* = h_{nb}(x) = \arg \max_{j=1, \dots, n} P(C_j|x) \prod_{i=1}^N P(X_i = x_i | C_j)$$

O canal que será seleccionado com uma maior probabilidade será uma função das co-ocorrências de determinadas combinações dos atributos em análise. Por exemplo, se em determinado dia for observado que Tipologia_SIC = "Telenovela", Tipologia_TVI = "Desporto", Tipologia_RTP1 = "Notícias", Syn_Ldin_1="V" e SynLdout_2="V" podemos calcular a probabilidade de selecção de cada um dos canais e atribuir como mais provável o que maximizar o valor calculado.

Realizando estas análises para cada elemento do painel ou para um colectivo, podemos calcular as taxas de cobertura mais prováveis de um programa. As taxas de cobertura representam o total dos telespectadores que contactaram pelo menos um segundo com determinado programa. Os *ratings* ponderam cada telespectador quanto à duração do contacto, logo uma vez atribuído a cada telespectador, ou a um colectivo, um canal será necessário a análise dos índices de fidelidade históricos. Considerando estas metodologias podemos tentar prever a selecção mais provável considerando os comportamentos de lealdade ao canal e programática, assim como as variáveis que caracterizam a tendência para se verificarem os efeitos de herança.

Os factores que condicionam uma selecção são inúmeros, mas neste contexto, existindo hábitos de consumo televisivo e uma exposição mais regular ao *media*, os comportamentos estruturais podem apresentar uma maior capacidade predictiva. O público no decorrer do tempo desenvolve comportamentos de lealdade a determinados programas, tipologias ou conjunto de canais, que em geral é restrito. Nem sempre, a contra-programação condiciona as selecções. Mas, ainda que existindo o hábito de ver determinado espaço informativo ou telenovela as selecções podem variar, como consequência das alterações que são realizadas na grelha de programas.

II. A construção dos modelos de exposição recorrendo aos modelos binomiais logísticos

Quando estão disponíveis os dados dos ficheiros base de audimetria, a selecção programática, também, poderá ser modelada por modelos *Binomiais Logísticos* (Chambers & Skinner, 2003). Pressupõe-se a regularidade da exposição ao *media* e que os segmentos se diferenciam quanto ao grau de lealdade que direccionam aos diversos suportes.

A análise *Logit Binomial* pode identificar quais são as características que influenciam a selecção e a probabilidade de um telespectador ser leal a determinado suporte.

Cada telespectador quando liga o televisor apresenta um conjunto de opções, sendo a probabilidade de seleccionar o programa i no instante t representada por p_t^i e a probabilidade de seleccionar um dos restantes programas dado pelo seu valor complementar, isto é, $1 - p_t^i$.

A selecção será função das características estruturais e das preferências dos telespectadores que podem ser representadas pela noção de utilidade. A probabilidade de seleccionar determinado programa será tanto maior, quanto mais os factores condicionantes da selecção se apresentarem como favoráveis. Logo, a lealdade programática pode ser modelada recorrendo a uma função Logística que prevê a ocorrência de um evento, e que neste contexto representa a selecção de um suporte. Para os níveis mais reduzidos da variável independente a probabilidade tende para zero, ou seja, para não se verificar a selecção. À medida que o valor das variáveis independentes cresce, a função tende para a unidade. Para valores de $p > 0,5$ prevê-se que ocorra a selecção.

Considerações teóricas e práticas sugerem que quando a variável resposta é binária, a forma da função resposta será, frequentemente, curvilínea. As funções logísticas, representam-se pela seguinte expressão:

$$E(Y) = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)}$$

Os coeficientes da Regressão Logística comparam a probabilidade de ocorrência e a probabilidade de não ocorrência, e representam-se pelo “odds ratio”. Os coeficientes com valores positivos levam a um acréscimo no “odds ratio”. A probabilidade de selecção por parte do telespectador i poderá ser representada por uma função da utilidade dada por:

$$P_i(\text{leal}) = \frac{\exp(V_i(\text{Leal}))}{\exp(V_i(\text{Leal})) + \exp(V_i(\text{não leal}))} \quad \text{onde } V_i(\text{Leal}) = \beta_0 + \sum_i \beta_i \text{Var}_i$$

As variáveis i representam os diversos factores que influenciam a selecção. Como factores que podem influenciar o grau de lealdade dos telespectadores podemos considerar as seguintes: as variáveis sociodemográficas (sexo, idade e classe social), as tipologias programáticas, variáveis associadas à lealdade ao *media*, tipológica ou programática, a estruturação da grelha, entre outras. Quanto mais estes factores se apresentarem como coadjuvantes, influenciando positivamente a selecção maior será a probabilidade de selecção, levando os telespectadores a situarem-se do lado direito da função Logística. A probabilidade de selecção e os factores que a condicionam relacionam-se segundo uma função em forma de “S”. A probabilidade de selecção será dada pela seguinte expressão:

$$\frac{e^{V_i(\text{Leal})} = \beta_0 + \sum_i \beta_i \text{Var}_i}{1 + e^{V_i(\text{Leal})} = \beta_0 + \sum_i \beta_i \text{Var}_i}$$

No conjunto dos factores a avaliar podemos considerar os seguintes:

➤ As variáveis sociodemográficas

Devem ser privilegiados os segmentos que apresentam valores mais variáveis dos indicadores de audimetria para cada um dos canais em análise. As variáveis sociodemográficas podem ser codificadas como variáveis binárias, sendo o seu valor função do segmento especificado a que cada telepectador corresponde.

Podem ser agregados os segmentos de telespectadores, que segundo as Análises Simbólicas, foram agrupados como apresentando evoluções mais semelhantes dos indicadores de audimetria. Devem ser privilegiadas as variáveis que discriminam em maior grau o público de cada um dos canais e para cada uma faixas horárias consideradas nas análises. Também, podemos considerar uma variável binária que traduza se o telespectador apresenta uma exposição televisiva acima da média para essas faixas horárias, apresentando essa variável o valor unitário se sim e nulo caso contrário.

Estes resultados podem ser consultados nas análises sociodemográficas que foram realizadas às taxas médias de audiências para o total da TV e aos *ratings*. Por exemplo, as variáveis podem ser representadas através das seguintes expressões:

$$\text{Faixa Etária} = \begin{cases} 1 & \text{se 4-24 anos} \\ 2 & \text{se 25-54 anos} \\ 3 & \text{se +55 anos} \end{cases} \quad \text{e} \quad \text{Classe Social} = \begin{cases} 0 & \text{se classe AB ou C1} \\ 1 & \text{se classe C2 ou D} \end{cases}$$

➤ Variáveis de caracterização do programa

Neste item podemos incluir as seguintes variáveis: 1. As tipologias programática de grau 1, grau 2 e grau 3 definidas no *software* Telereport; 2. A semana em que se transmite o programa desde o seu início. A transmissão dos episódios iniciais de um programa não apresentará uma exposição equivalente ao que se verifica com um programa que se encontre num estágio mais avançado do seu ciclo de vida; 3. Se o programa consiste numa retransmissão, por exemplo, codificando-o como nulo para o programa original e em um para uma eventual retransmissão.

➤ Variáveis que caracterizam a interacção do telespectador com o meio televisivo, por exemplo, as seguintes variáveis de posicionamento programático.

A posição do programa poderá ser representada por quatro variáveis binárias representadas pelas seguintes expressões:

$$X_{\text{Tv:On-Ldout}} = \begin{cases} 1 & \text{O telespectador assistiu a outro canal nos} \\ & \text{30 minutos posteriores} \\ 0 & \text{Caso contrário} \end{cases}$$

$$X_{\text{Tv:On-Ldin}} = \begin{cases} 1 & \text{O telespectador assistiu a outro canal nos} \\ & \text{30 minutos anteriores} \\ 0 & \text{Caso contrário} \end{cases}$$

$$X_{\text{ldout}} = \begin{cases} 1 & \text{O telespectador assistiu ao mesmo canal nos} \\ & \text{30 minutos posteriores} \\ 0 & \text{Caso contrário} \end{cases}$$

$$X_{\text{Ldin}} = \begin{cases} 1 & \text{O telespectador assistiu ao mesmo canal nos} \\ & \text{30 minutos anteriores} \\ 0 & \text{Caso contrário} \end{cases}$$

Como variável dependente considerar se o telespectador assiste ao programa p no instante t e que será representado por uma variável binária:

$$y = \begin{cases} 1 & \text{telespectador assiste programa} \\ & p \text{ no instante } t \\ 0 & \text{telespectador não assiste} \end{cases}$$

III. A construção de modelos de exposição a séries de TV por análises dos dados de fluxos de audiências

Segundo Wyndham (1994), as audiências dos diversos episódios de uma série dependem dos seguintes factores:

- (1) O número de telespectadores que assistiu ao episódio X e, também, assistiu ao anterior;
- (2) O número de telespectadores que se repetem com determinado intervalo de tempo;
- (3) O número de telespectadores que nunca assistiram a nenhum episódio anterior a X .

Para promover a previsão das audiências de séries de TV, o autor refere ser necessário avaliar de que depende cada uma das componentes que constitui a sua audiência e, na sua sequência constatar se e como podem ser previstas.

Ou seja, devemos compreender quais são os factores que condicionam a repetição da exposição dos telespectadores a dois episódios de um programa, que podem ser consecutivos ou não, e o que determina a sua exposição inicial e que poderá não se verificar novamente no futuro ou verificar-se de forma muito esporádica.

Em *Wyndham (1994)*, descreve-se uma abordagem que tenta viabilizar a previsão das audiências totais de uma série que deverá apresentar entre 6 a 12 episódios. Por conseguinte, inviabiliza a sua aplicação quando se pretende quantificar as audiências de telenovelas ou de séries com uma maior duração, como se verifica, frequentemente, a nível nacional. Apesar do reduzido número de suportes considerados, o autor reconhece as limitações deste modelo. Apresenta erros de previsão no valor de 30% quando recorre aos dados que caracterizam apenas o primeiro episódio e no valor de 10% quando se considera alguma informação a respeito dos dois episódios iniciais de uma série.

Na presente tese foram realizadas algumas análises, que nos permitiram verificar que os índices de duplicação entre episódios adjacentes ou segundo um maior afastamento temporal, ainda que apresentando alguma variabilidade, nem sempre o seu valor é muito significativo. Portanto, por vezes podemos assumir como sendo relativamente constante o seu valor, mas nem sempre, como o faz este autor no modelo apresentado.

Wyndham define o êxito de uma série, comparando o valor médio da audiência de todos os episódios com o valor da audiência dos dois episódios iniciais. No modelo desenvolvido considera a seguinte informação: os índices dos dois episódios iniciais, o índice de duplicação entre os dois primeiros episódios, um índice de apreciação (IA) que corresponde a quem “gostou” do programa, um índice de concorrência que será o valor do IA do primeiro episódio comparado com a média dos IA de todos os restantes episódios e um índice que resulta da divisão entre o valor da cobertura de quem assiste pelo menos 3 minutos do episódio e de quem assiste apenas um minuto.

Na segunda questão em avaliação, na presente tese, demonstrou-se que as selecções não resultam, frequentemente, de preferências, existindo um vasto conjunto de factores estruturais e individuais que podem determinar a exposição. Consequentemente, é insuficiente tentar promover a previsão recorrendo a uma medida de quantificação da preferência.

De acordo com as análises de duplicação que foram realizadas, no máximo apenas metade da audiência de um programa é transportada para os seus sucessivos episódios. A disponibilidade e a predisposição para se verificar a exposição ao *media* condiciona a dimensão da duplicação das audiências, sendo os valores programáticos de duplicação semelhantes aos que resultam de considerar os valores equivalentes para as faixas horárias. Consequentemente, tentar quantificar apenas a duplicação das audiências, também, será insuficiente para a construção dos modelos de exposição. A dimensão da influencia da irregularidade da exposição ao *media* pode superar a repercussão da irregularidade observada ao nível da selecção dos conteúdos.

Para cada suporte considerado, uma percentagem da audiência será herdada de episódios anteriores, ou seja, dos comportamentos de lealdade programática, que, também, podem apresentar alguma ciclicidade quando se consideram eventos com algum afastamento temporal. Para esta fracção do painel, em função das suas selecções passadas podemos calcular uma probabilidade de selecção e, posteriormente tentar calcular a duração mais provável ou representativa da sua exposição. Devemos considerar um longo horizonte temporal, uma vez que os comportamentos não apresentam ciclos com uma duração reduzida.

Porém, para uma elevada percentagem do painel, que não apresenta uma exposição mais regular ao meio televisivo, será muito difícil conseguir prever quando efectivamente irá ligar o televisor e as respectivas selecções, inviabilizando esta perspectiva de análise.

Se nos restringirmos aos segmentos que apresentam uma exposição mais regular, será necessário quantificar a duplicação de audiências entre os episódios, mas, também, uma variável relativa aos efeitos de herança, que por sua vez se podem interseccionar ou não, ou resultarem dos fluxos de audiências por parte de canais concorrentes.

A audiência de um programa encontra-se distribuída por diferentes componentes e representadas nos seguintes itens:

- (1) $X_1\%$ da audiência do mesmo programa transmitido no dia anterior ou em dias anteriores.

O programa pode variar nos horários de transmissão, inculcando variações na evolução destes indicadores. Será função do grau de lealdade programática, assim como dos factores de disponibilidade e vontade de ver TV e da tipologia programática em análise. O conteúdo programático, também, poderá influenciar os resultados, e, logo, os episódios que apresentem cenas que possam suscitar um maior interesse dos telespectadores podem apresentar maiores valores para estes índices. A repercussão do conteúdo programático será dificilmente quantificável, como se demonstrou para os eventos desportivos. Atendendo às análises realizadas, verificamos que existe um segmento que recorre ao meio televisivo, essencialmente, para assistir aos espaços informativos ou a uma telenovela. Nestes contextos as análises de lealdade programática será mais relevante para prever os valores de *ratings*, em função da sua probabilidade de selecção.

Para os telespectadores que apresentam uma exposição mais regular, podemos calcular uma probabilidade de selecção e a duração de exposição média recorrendo aos dados históricos. Alguns destes telespectadores optam por desligar o televisor após a transmissão do programa, enquanto outros são transportados por um efeito de inércia para os programas subsequentes. Estamos perante os comportamentos de lealdade programática, eventualmente conciliada com os efeitos de herança. As análises de lealdade programática também podem incidir nas faixas horárias sucessoras e verificarmos qual o grau de duplicação de audiências.

(2) $X_2\%$ da audiência do programa de “Lead_In” ($y\%$ TV_OFF, $z\%$ muda de canal).

Este valor será função do grau de passividade das selecções, factores de disponibilidade, lealdade ao canal e tipológica e da programação concorrente. Por exemplo, face à transmissão de uma telenovela no canal SIC, a audiência poderá manter-se neste canal assistindo ao programa subsequente ou seleccionar uma telenovela que considerem mais interessante da TVI se existir lealdade tipológica ou um filme da RTP1 se existir uma procura por complementaridade programática.

Será necessário avaliar quais são os comportamentos estruturais que caracterizam em maior grau cada segmento de telespectadores. Nos períodos que apresentem uma maior regularidade nos padrões de disponibilidade será superior a probabilidade de seleccionarem um programa que apresente uma sequência narrativa, podendo seleccionar em faixas horárias mais tardias programas cujos episódios não apresentem interdependências face a programas anteriores. No entanto, atendendo às análises realizadas, a duplicação de audiências encontra-se essencialmente influenciada pelas decisões de ligar ou desligar o televisor e a associação existente entre a duplicação das audiências e os *ratings* é muito significativa. Logo, a variável mais relevante para a previsão dos *ratings* corresponde à previsão dos factores de disponibilidade do painel e de “vontade” para ver TV que dificilmente será quantificável e que atendendo às análises realizadas aos índices de lealdade programática apresenta uma considerável variabilidade.

(3) $X_3\%$ da audiência do programa de “Lead_Out” ($y\%$ TV_OFF, $z\%$ assiste outro canal).

Ligar o televisor em períodos anteriores ao programa que os telespectadores pretendem assistir será função da sua disponibilidade. A associação dos valores de *ratings* face aos valores sucessores é muito menor comparativamente aos valores antecessores, segundo as análises de duplicação realizadas.

(4) $X_4\%$ da audiência que apresenta o televisor ligado nas faixas horárias de “Lead_In”, mas vendo um canal concorrente ($y\%$ TV_OFF e $z\%$ não muda de canal ou muda para um canal diferente do programa em análise). Este valor será uma função da lealdade tipológica e programática.

(5) $X_5\%$ da audiência que apresenta o televisor ligado nas faixas horárias sucessoras, mas assistindo a um canal concorrente ($y\%$ TV_OFF e $z\%$ não muda de canal ou muda para um canal diferente do programa em análise). Este valor será função da lealdade tipológica e programática e dos factores de disponibilidade.

As questões associadas a $X_4\%$ e $X_5\%$ podem ser as menos relevantes para a previsão dos *ratings*.

(6) $X_6\%$ da audiência liga o televisor no início desse programa. Este valor será função da lealdade programática condicionada pelos factores de disponibilidade. Nestes contextos o valor terá que ser quantificado em função dos índices de lealdade programática. As análises podem ser conduzidas em função das faixas horárias e, posteriormente, será necessário extrapolar para programas em concreto, considerando as durações parciais e as globais.

Posteriormente, será necessário calcular durante quanto tempo se verificará a exposição para proceder ao cálculo das taxas médias de audiências que ponderam os telespectadores em função da duração da exposição. Também, têm que ser consideradas as alterações que se verificam nas componentes anteriores quando são transmitidos os eventos desportivos, os quais apresentam repercussões muito diferenciadas.

Relativamente à previsão de cada uma das componentes descritas anteriormente, por exemplo, para a quantificação da duplicação de audiências entre sucessivas transmissões de um programa podemos recorrer à seguinte sequência de análises:

A percentagem da audiência herdada do período anterior, em média, será $P(t_3 \wedge t_2) = P(t_3|t_2) * P(t_2)$. $P(t_3|t_2)$ será equivalente à duplicação de audiência média entre t_2 e t_3 e em percentagem. Coloca-se a questão de qual é a duração da exposição a considerar nas análises. Se for independentemente da duração do contacto a duplicação de audiências terá que ser calculada seleccionando um segundo.

$P(t_2)$ será o valor da taxa de cobertura do programa que não pondera os telespectadores relativamente à duração da exposição, uma vez que mede a percentagem de indivíduos que contactam com o evento pelo menos um segundo. $P(t_3 \wedge t_2)$ será quem assiste a ambos os programas independentemente da duração da exposição. As taxas médias de audiências consideram o total de segundos contactados, logo face a $P(t_3 \wedge t_2)$ teríamos que multiplicar pelo valor de Ats associado ao período t_3 (tempo médio que cada telespectador despende a ver um evento).

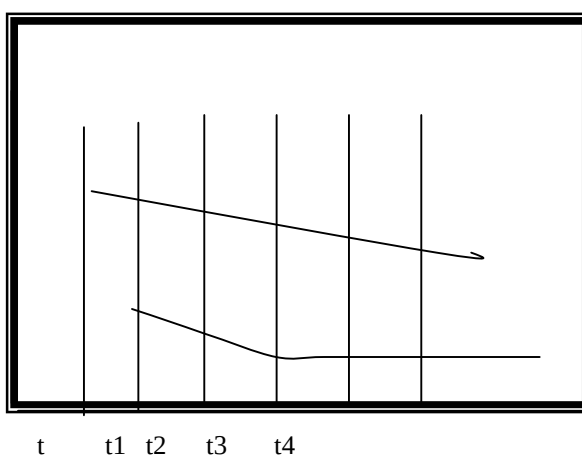


Fig. VI.4.4.1.2 – Duplicação de audiências entre períodos horários.

Fonte: Elaborado pela autora.

A melhor estimativa do contributo da audiência do período anterior para o cálculo das taxas médias de audiências, seria $(*Ats(t_3)) / (Duração\ de\ t_3 * n.^o\ de\ indivíduos\ do\ alvo)$.

Podemos considerar como duração os 100%, e, logo será quem assiste a t_3 na totalidade, sendo o seu contributo para a duração de segundos contactados $P(t_3 \wedge t_2) * \Delta t_3$.

Podemos considerar os 50% da duração de t_3 , sendo o seu contributo para a duração de segundos contactados representada por $P(t_3 \wedge t_2) * (\Delta t_3) * (1/2)$.

Percentagem da audiência herdada de t_{+1} em média, independentemente do tempo de contacto. Só devemos contabilizar os telespectadores que seleccionam t_{+1} e t_{+3} , mas excluindo a duplicação de audiências entre t_{+2} e t_{+1} , uma vez que a duplicação de audiências face a t_{+2} já foi calculada.

$$P(t_3|t_1) = \frac{P(t_3 \wedge t_1)}{P(t_1)}, \text{ mas interessa excluir de } t_{+1} \text{ quem viu } t_2.$$

Audiência significativa herdada de $t_1 = P(t_3|t_1) * p(\bar{t}_2 \wedge t_1) = P(t_3|t_1) * (P(\bar{t}_2 / t_1) * P(t_1))$

$P(t_3|t_1)$ será o valor médio observado de duplicação entre t_3 e t_1 ; $(P(\bar{t}_2 / t_1))$ será o índice de exclusividade médio entre t_2 e t_1 e $P(t_1)$ será a taxa de cobertura média para t_1 .

Para o cálculo da audiência média uma vez mais seria necessário recorrer ao valor de Ats. $P(t_3|t_1)_{efectiva} * \text{Ats}(t_3) / (\text{Duração de } t_3 * n.^{\circ} \text{ de indivíduos do alvo}).$

O mesmo procedimento teria que ser aplicado aos restantes períodos que antecedem o período de referência. Esta abordagem destina-se ao cálculo de um valor médio sem considerar a influência dos factores estruturais. O total de segundos contactados seriam os que decorrem da duplicação de audiências face a cada um dos Δt anteriores. Posteriormente, os valores associados a programas teriam que ser calculados, em função dos horizontes temporais que abrangem. Perante uma grelha em que os programas apresentam grandes oscilações horárias e durações, as análises de duplicação de audiências conduzidas com base em valores programáticos apresentam uma forte variabilidade, não sendo viável a sua previsão. Em 2010, os canais apresentam uma maior consistência nas suas transmissões e uma considerável estabilidade horária. Nestes contextos será mais fácil definir um intervalo de variabilidade para os índices de duplicação associados a cada um dos programas e avaliar a repercussão dos factores estruturais. Todavia, a audiência de um programa ou período, também, será função das restantes componentes da audiência que apenas poderá ser avaliada recorrendo a determinadas metodologias de *Data Mining* e que se descreve nos seguintes capítulos.

VI.4.4.2. Outros contextos de construção de modelos de exposição quando existe regularidade na exposição

A análise dos comportamentos de lealdade e a análise do padrão de duplicação de audiências é fundamental para os segmentos II e III. Contudo, a construção dos modelos de exposição por quantificação da duplicação das audiências não será, certamente, a melhor perspectiva de análise e devem ser testadas abordagens alternativas. A interacção que possa existir entre os diferentes comportamentos estruturais e os factores de planeamento condicionará em maior grau a evolução dos valores de *ratings*. Uma vez que partimos do pressuposto destes segmentos poderem apresentar uma maior regularidade do consumo televisivo, a identificação das Regras de Associação, que não têm em atenção a sequência temporal subjacente às selecções, ou as Regras Sequenciais mais características será, eventualmente, a melhor perspectiva para conseguir identificar os conteúdos a contratar em cada um dos subsegmentos existentes. Cada telespectador apresentará uma sequência de exposição, que até poderá variar de dia para dia, e, por vezes, poderá existir total ausência de exposição ao meio televisivo, porém o mais relevante será identificar o seu comportamento mais habitual e característico.

Para o horizonte temporal em análise verificamos que no segmento II existem diferentes durações de exposição. Uma fracção desliga o televisor após a principal telenovela da noite e o público que permanece por um maior horizonte temporal, subdivide-se no público que permanece no mesmo canal no decorrer de toda a noite e no público que transita entre canais, em especial entre os dois canais privados generalistas. No segmento III, verifica-se a mesma situação, mas deverá existir uma maior lealdade ao canal e, consequentemente, uma maior duplicação de audiências no decorrer do dia e para o mesmo canal. Nestes contextos, será necessário recorrer aos dados dos ficheiros base de audimetria e as abordagens metodológicas mais adequadas procedem do *Data Mining* Temporal, como, por exemplo, os diversos algoritmos de Regras de Associação e Sequenciais e correspondentes algoritmos de *clustering*. Com base nestas metodologias, podemos tentar identificar quais são as selecções programáticas mais prováveis de ocorrer em determinados horizontes temporais e segmentar a base de dados em função deste factor.

Correspondem a proposições probabilísticas sobre a ocorrência de determinados estados. As Regras podem considerar atributos com vários valores, serem multidimensionais ou multinível. Cada Regra $A \rightarrow B$ representa uma sequência de selecções e é caracterizada pela sua frequência relativa, ou seja, pela percentagem de telespectadores onde se verifica, assim como pelo número de sessões televisivas que satisfazem B, condicionadas pela presença de A, representando a probabilidade de selecção de B se A se verificou. Nas Regras de Associação, também, existe a necessidade de considerar um “trade-off” entre o valor da frequência e da cobertura alcançada pelo conjunto de suportes e recorrendo ao menor investimento possível.

Uma vez definidas as sequências mais características de cada telespectador e, posteriormente, para cada segmento, os conteúdos a considerar devem apresentar uma considerável frequência relativa, ou seja, permite avaliar a cobertura do alvo e do número de sessões de cada telespectador que integra o painel, assim como a frequência condicionada, uma vez que se o seu valor for reduzido significará que a exposição será em maior grau aleatória e não existirá uma complementaridade estrutural dos conteúdos que incluem. Dividindo a frequência de cada regra e o valor da sua probabilidade condicional obtemos um valor que representa o número de vezes que a probabilidade de estar exposto a B condicionada pela presença de A na sessão supera uma exposição aleatória. As sequências podem ser directas ou indirectas, representando o diferencial de frequências para ambos os contextos as ocasiões em que os suportes alternativos separam os que se encontram em avaliação.

No planeamento de *media* pretende-se avaliar a frequência de contacto individual e do segmento face ao conjunto de suportes, ou seja, x% contacta com um suporte, y% contacta com 2 suportes, e assim sucessivamente. Considerando a totalidade das sequências de um telespectador no decorrer de um grande horizonte temporal e que será viável com a audimetria, as suas frequências relativas correspondem ao número de vezes que estas ocorrem no global de sequências em análise.

Logo, uma sequência de ordem 2 que apresenta uma frequência de 40%, significa que esses dois suportes apresentam uma probabilidade de contactar com 40% do painel ou com um telespectador no decorrer do período associado à campanha. Uma sequência de ordem três que apresente uma frequência de 20% permite-nos concluir que 20% do painel poderá contactar com esses três suportes se estes forem seleccionados. As frequências condicionadas representam o número de vezes que existiu contacto com determinado suporte, sendo verdadeiro o seu antecedente. Portanto, se uma sequência de grau três integrar duas sequências de grau dois e apresentar probabilidades condicionadas de 20% ou 40% deverá ser seleccionada a que apresenta uma probabilidade de 40%, uma vez que a duplicação das audiências entre os suportes será superior, a não ser que o diferencial ao nível dos custos não justifique optar por uma maior probabilidade condicionada.

Com a utilização das Árvores podemos identificar quais são as sequências com uma maior probabilidade condicionada, seleccionando como variável alvo a que apresenta uma maior cobertura e os suportes que nas diversas ramificações até aos nós terminais apresentam todos o valor unitário, isto é, que estão presentes. As sequências apresentam diversas ordens de grandeza, sendo mais relevantes as que conseguem alcançar uma considerável fracção do público-alvo, com um número mínimo de suportes e que devem ser seleccionados em função dos custos que lhe estão associados, mas não só. Poderá existir um maior interesse em seleccionar suportes de determinado segmento horário ou canal ou tipologia. Podemos avaliar quais são os suportes associados em maior grau às variáveis TV_ON e a TV_OFF.

Posteriormente, o *Link Mining* (Getoor, 2005) viabiliza a visualização dos suportes que apresentam uma maior ligação ou ausência de ligação, ou seja, os suportes que são seleccionados conjuntamente em maior grau numa determinada sequência temporal. Considerando as contagens para cada par de suportes e por análise MDS (Wickelmaier, 2003) dos dados, podemos visualizar as proximidades num espaço eucladiano bidimensional. Quanto maior for a duplicação entre eles mais será a proximidade entre os suportes.

Consequentemente, uma vez identificados os suportes que apresentam uma maior exposição no seu conjunto, poderá ser avaliada a duplicação das audiências entre eles recorrendo às frequências condicionadas, ao *Link Mining* e às análises MDS. A selecção final será função das características da campanha em avaliação e da aplicação de algoritmos de optimização.

Quando são considerados suportes programáticos, em vez de suportes horários, a duração da exposição terá que ser avaliada com base no histórico de valores e, preferencialmente, recorrendo a variáveis simbólicas. A análise dos indicadores históricos de fidelidade é essencial, uma vez que quanto maior for o grau de envolvimento com os suportes, maior será o desempenho das inserções comerciais que lhe estão associadas. As Regras, também, podem ser calculadas considerando a exposição a faixas horárias. Uma vez que neste contexto aceitamos a existência de hábitos de consumo televisivo, deverá existir uma maior relação dos telespectadores face ao *media*, e não, perante os conteúdos em concreto e, por conseguinte, as análises devem ser conduzidas por suportes horários. Foi possível avaliar a grande influência da estrutura horária na evolução dos principais indicadores de audimetria.

Tradicionalmente para estes segmentos considerava-se a frequência de exposição, que na realidade não é válida. A concepção clássica de probabilidade apenas se verifica em grandes horizontes temporais, mas, não, é possível demonstrar, ou seja, uma probabilidade de 4/7 não significa que o comportamento da audiência se ajuste exactamente a essa probabilidade numa semana concreta. Mediante a observação e a experimentação não é viável conhecer as probabilidades de realização de determinado comportamento.

Para os telespectadores que apresentam uma maior regularidade de exposição e uma significativa duplicação de audiências, as alterações no seu padrão de consumo televisivo em princípio devem-se a acontecimentos inesperados. Se um telespectador vê determinado programa em dias concretos deverá continuar a fazê-lo, a não ser que surjam imprevistos que o impossibilitem.

Nestes contextos os comportamentos são mais estáveis e o telespectador eventualmente poderá escolher o que mais gosta, ou em maior grau o que menos lhe desagrada, mas o seu comportamento habitual será estar frente ao televisor nessas faixas horárias. Quando os programas se repetem cada semana o telespectador não terá que realizar nenhuma decisão se tem o hábito de ver as sucessivas emissões.

Alternativamente as análises de médias móveis pesadas, os algoritmos Bayesianos e as metodologias Logit Binomiais, também, permitem avaliar os comportamentos de lealdade e a probabilidade de selecção do *media* ou de determinado programa. As análises Logit Binomiais permitem avaliar os factores que determinam a selecção, mas, não, comportamentos concretos no espaço e tempo, e, logo, se existir irregularidade, são, por si só, insuficientes. Porém, neste contexto assume-se a existência de hábitos de consumo televisivo e de uma maior regularidade da exposição. Se considerarmos os dados agregados por segmentos existem inúmeras abordagens metodológicas procedentes, quer do *Data Mining*, quer da Análise Simbólica. As Regras Sequenciais e de Associação podem ser aplicadas a segmentos, desde que estes apresentem padrões equivalentes de exposição televisiva.

Com as Regras de Associação facilmente podemos quantificar a cobertura do alvo considerando os dados históricos e que, também, será o seu valor mais provável para o curto prazo se esses suportes forem seleccionados, sendo o seu valor equivalente ao da frequência de uma Regra. Esses suportes correspondem aos que, no passado, foram seleccionados conjuntamente com uma maior frequência, sendo as suas subsequências, também, as mais frequentes. Consequentemente, em relação ao valor da frequência será a maior e mais característica para esse segmento. Tradicionalmente os planos, também, são avaliados em função da sua distribuição de frequências, que, também, pode ser facilmente calculada para o histórico e, neste contexto, existe uma maior regularidade da exposição. Identificadas as sequências levanta-se a questão de quais são os conteúdos a incluir nos planos.

Devem maximizar a cobertura e para um valor equivalente deste indicador em função das características da campanha e do número mínimo de exposições mais adequado, podemos seleccionar os conteúdos atendendo à distribuição de frequências. Por exemplo, se em determinada sequência X, a percentagem do alvo que contacta três ou mais vezes com esses conteúdos for superior face às restantes sequências, devemos privilegiar esses conteúdos. Também, podemos conhecer para o histórico qual foi o grau de duplicação das audiências, o perfil do público e a sua fidelidade aos conteúdos. Os algoritmos de optimização permitem considerar as restrições orçamentais, mas como ficou previamente demonstrado, é necessário proceder a uma pré-selecção dos suportes.

VI.4.5. Os modelos de exposição em contextos de inexistência de hábitos de consumo televisivo e numa óptica de construção de planos de *Media*

Quando estão em análise comportamentos de consumo televisivo mais irregulares, por exemplo, para o segmento IV e VI descritos na *Figura VI.4.2.1* será muito mais difícil construir um modelo de exposição. A audiência poderá ser pouco consistente no decorrer do tempo, não estando consolidados os hábitos de ver TV e as selecções dependem, essencialmente, do modo de vida dos telespectadores. Nestes segmentos incluiríamos os “zappers” que não planificam o seu consumo televisivo e que em função da sua disponibilidade e da predisposição para assistir à TV transitam de canal para canal na procura da melhor alternativa disponível. Em geral não assistem aos programas na integra, contribuindo para uma maior variabilidade dos índices de audiências.

Em primeiro lugar será necessário considerar o colectivo, uma vez que o comportamento individual é imprevisível. Verificamos que a audiência de um programa ou faixa horária nestes contextos será difusa. A audiência de um programa evolui no tempo e a audimetria assim o demonstra, apresentando, por vezes, uma forte irregularidade que poderá ser atribuída em maior grau a estes segmentos.

Cada indivíduo apresenta uma tendência a fazer parte da audiência de determinado programa e que não depende, certamente, apenas dos conteúdos programáticos transmitidos, mas, também, dos factores estruturais, entre os quais se destacam a disponibilidade e a estrutura do comportamento individual relativamente à totalidade dos programas e ao meio televisivo, assim como as rotinas diárias que este desenvolve no seu quotidiano.

Neste contexto, não se verificam comportamentos habituais com respeito aos programas, faixas horárias e muito menos em relação aos possíveis conteúdos dos programas, o que dificulta as análises de previsão. Correspondem a comportamentos não habituais para uma fracção do painel, mas de alguma forma, talvez, repetitivos em relação a determinado conjunto de conteúdos de programas.

O telespectador poderá apresentar um conjunto de programas que irá seleccionando no decorrer do tempo e que poderá variar, assim como com uma frequência variável. A nível estatístico a previsão nestes contextos terá que resultar da análise dos dados para o colectivo, na esperança de que algumas regularidades possam ser identificadas.

O fenómeno social não depende da natureza dos indivíduos. É que, na fusão de que ele resulta, todos os caracteres individuais, divergentes por definição, se neutralizam e se dissipam mutuamente. Só as propriedades mais genéricas da natureza humana sobrevivem. Devido precisamente à sua extrema generalidade, é que elas nunca poderiam justificar as formas muito especiais e complexas que caracterizam os factos colectivos. Não é que eles tenham nada a ver com o resultado; mas isso apenas se verifica em condições imediatas e longínquas. O resultado não se produziria se elas o excluíssem; mas não são elas que determinam esse mesmo resultado. (*Durkheim, 1982, p.198*)

A integração das consciências individuais vai resultar em algo que é composto pela síntese do todo, que tem força de ultrapassar a parte, ou seja, o indivíduo. ... os fenómenos sociais não dependem da natureza individual, ambos possuem leis próprias. ... o social apresenta uma realidade de natureza diversa da do indivíduo, não constituído apenas pela soma destes, mas pela sua associação. (*Lima, 2007, p.8*)

Nestas situações será, talvez, mais relevante identificar quais são os diversos conteúdos que procuram, independentemente da regularidade e duração da exposição aos mesmos.

As abordagens metodológicas podem em princípio transcender as análises quantitativas e ser complementadas com as análises qualitativas.

Quando existe uma maior irregularidade da exposição, na prática é frequente recorrer às sondagens de intenção de ver em que cada telespectador atribui uma nota a cada programa em função do interesse que lhe desperta e a partir desses dados realizam-se as estimativas. Porém, está internacionalmente demonstrado que o interesse não corresponde a uma variável, por si só, explicativa do consumo televisivo e ainda menos de previsão e, provavelmente, ainda menos para este segmento que apresenta um reduzido envolvimento com o *media* e os conteúdos. As abordagens qualitativas alternativas, também, apresentam grandes limitações. A dimensão da repercussão da contra-programação será função do grau da predisposição para se desenvolverem comportamentos estruturais de consumo televisivo.

Consequentemente, a análise dos fluxos de audiências entre canais apresenta uma maior relevância, assim como a utilização de abordagens simbólicas que permite uma representação dos dados que traduza a variabilidade e instabilidade observada nos dados. Algumas tipologias de programas, também, apresentam variações no seu desempenho, e, não, como consequência das alterações observadas nos seus contextos de contra-programação, mas, sim, devido ao efeito “sobredose”, ou seja, se os telespectadores já assistiram anteriormente a vários filmes iguais ou que possam ser considerados semelhantes. Dificilmente as análises de previsão qualitativas conseguem ter este aspecto em consideração.

As sondagens de intenção de ver, também, apresentam inúmeras dificuldades práticas como, por exemplo, não se dispor da listagem de programas dos canais e, por vezes, quando se consegue não será com a necessária antecedência para se poder realizar a consulta ao painel e preparar a previsão. A excessiva concorrência entre os canais, por vezes, torna impossível dispor de listas exactas de programas que se vão emitir com a necessária antecedência. Estes sistemas permitiam que melhores resultados fossem alcançados, quando existia uma menor rivalidade entre as estações televisivas.

Os inquéritos, também, não têm em atenção as variações na programação de último minuto, assim como, as transmissões relacionadas com a actualidade, como, por exemplo, o desaparecimento de alguém e não se adequam a contextos de maior irregularidade dos dados.

Em princípio, deverá ser diferente tentar prever as audiências após o horário nobre, uma vez que os canais podem transmitir em diferentes horizontes temporais. Quando um programa conclui, os telespectadores podem não ter uma grande opção de escolha de continuidade, ao contrário, do que se verifica no início do espaço informativo. Em especial para a análise dos segmentos IV e VI, ainda que, também, poderá ser aplicada em outros contextos, apresenta-se a seguinte sequência de análises:

- Análise da exposição global a um conjunto de suportes (ETAPA I);
- Análise de manifestações mais concretas da exposição em determinada semana e perante os suportes seleccionados para integrar os planos de *media*. Na realidade serão análises probabilísticas, sendo necessário seleccionar os conteúdos que apresentam uma maior propensão para serem seleccionados nesse horizonte temporal (ETAPA II).

O segmento IV apresentará uma maior variabilidade ao nível da selecção e o segmento VI, também, ao nível da exposição ao *media*, sendo mais difícil, neste contexto, realizar as análises de previsão. As Ciências Sociais são classificadas como inexactas, e, logo, as previsões apresentam características que as diferenciam das análises que incidem nos sistemas físicos.

Em *Bunge (1981)*, salienta-se que a Ciência Social beneficia de um estudo teórico e empírico mais intenso das probabilidades de transição, para, assim, poder prever a direcção da evolução mais provável dos acontecimentos. Segundo o autor, isto é, o que se faz em física atómica e na vida quotidiana, com a particularidade que neste último caso a projecção estocástica é geralmente qualitativa e não quantitativa e sem aprofundar ou debilmente fundamentada e sem serem baseadas em teorias científicas.

O autor considera que a projecção estocástica é cientificamente respeitável e que devemos ter isto presente se se pretender uma maior aceleração da evolução das ciências não físicas. Salaria que a projecção estocástica, colectiva (estatística) ou individual (probabilística) não deve ser considerada como menos importante e interessante do que a projecção não estocástica (“determinista”), em especial se for realizada com a ajuda de teorias e sobre a base de generalizações estatísticas empíricas.

A seguinte *Figura VI.4.5.1* representa a sequência de análises que é proposta nestes contextos. A construção de um plano de *media* deverá abranger uma sequência de etapas que dependem dos objectivos de determinada campanha publicitária, sendo as mais relevantes as seguintes:

- (1) Segmentar a emissão diária em função da variabilidade dos indicadores e conduzir as análises por segmento horário identificado (poderá ser facultativo);
- (2) Escolha do segmento em função das variáveis que caracterizam o público-alvo da campanha publicitária e que podem ser complementadas com os resultados das análises de segmentação do painel de audimetria em função dos respectivos perfis de exposição;
- (3) Identificação do conjunto de programas ou suportes horários a que os telespectadores seleccionados se encontram expostos, independentemente da frequência e regularidade da exposição, e construção das sequências “mais características”, diferenciando nomeadamente os dias laborais dos dias de fim-de-semana, assim como os períodos de férias;
- (4) Identificação de segmentos através do *clustering* das sequências de exposição mais características de cada telespectador;
- (5) Análises de caracterização de cada segmento e a “previsão da exposição global” através da identificação de Regras Sequenciais ou de Associação ou de Dependência mais representativas para cada segmento previamente identificado.

Com as Regras de Associação podemos identificar os contextos horários mais adequados e com as Regras Sequenciais podemos identificar os conteúdos que se podem repartir por diferentes *clusters* horários;

(6) Complementar a selecção de suportes em função dos resultados da associação de conteúdos aos públicos-objectivos e às variáveis de caracterização do painel de audimetria recorrendo aos métodos factoriais e dos resultados das análises de duplicação das audiências. Podemos recorrer ao *Link Mining* e ao MDS para visualizar as associações e avaliar o grau de duplicação das audiências entre os suportes identificados nas Regras;

(7) Análises de previsão para uma semana em concreto recorrendo aos modelos explicativos de exposição televisiva e ao *Data Mining* não proposicional. As análises devem ser realizadas face aos suportes previamente identificados como pertencendo às sequências mais características e por análise das associações dos conteúdos às variáveis de caracterização do painel.

A redução da incerteza dos resultados será mais difícil de alcançar para o segmento VI que apresenta irregularidade na exposição ao *media* como resultado da sua predisposição e vontade para ver TV ou das suas vivências diárias. Quando impera uma maior irregularidade da selecção será mais difícil identificar variáveis que permitam a construção de Regras de classificação ou de previsão;

(8) A selecção dos suportes comerciais em função da evolução mais provável das audiências no decorrer de um espaço comercial que é função das variáveis de planeamento que os caracterizam no seu todo e das diversas inserções;

(9) A construção do plano final com algoritmos de optimização, em função das restrições orçamentais, entre outras, e considerando o conjunto de suportes que foram seleccionados.

Devemos ter em atenção os suportes que foram identificados nas etapas anteriores, devendo-se privilegiar uma maior cobertura do segmento e podendo-se atribuir uma maior ponderação aos telespectadores e segmentos que apresentam uma maior semelhança face ao perfil do público-alvo da campanha publicitária. Devemos adoptar como frequência de exposição a que for mais adequada ao produto, “marca” e segmento;

(10) Necessidade de reavaliação dos resultados no decorrer do tempo, recorrendo a Regras Sequenciais para ambientes dinâmicos (*Amornchewin & Kreesuradej, 2009*). As Regras que são actualizadas em função do seu grau de dissemelhança face aos resultados históricos e segue-se um reajustamento dos restantes resultados que delas decorrem.

VI.4.5.1. A análise da exposição global

Nos seguintes parágrafos, descreve-se algumas das etapas que forma mencionadas anteriormente. Sendo os comportamentos individuais imprevisíveis, e atendendo ao elevado número de suportes que podem vir a ser incluídos nos planos de *media*, será necessário restringir o seu número e incidir as análises nos conteúdos e nas faixas horárias que se associam em maior grau com os telespectadores que integram o público-alvo pretendido. Identificando o conjunto de suportes a que cada segmento se encontra mais exposto irá permitir-nos prever a exposição a nível global, independentemente da variabilidade que se observa entre cada semana e dia considerado.

I. Segmentar a emissão diária em função da variabilidade dos indicadores e conduzir as análises por segmento horário

Na primeira etapa das análises, se necessário proceder em primeiro lugar à segmentação da emissão diária que deverá ser função da variabilidade observada nos valores de *ratings*.



Fig. VI.4.5.1 – Etapas que viabilizam a identificação dos suportes publicitários nos quais investir.
Fonte: Elaborado pela autora.

Conduzir as análises nas etapas posteriores por segmento horário identificado, recorrendo a metodologias simbólicas ou de *Data Mining* Temporal. Estas análises são essenciais do ponto de vista da definição da amplitude horária dos suportes a considerar. Nos segmentos horários onde se observa uma maior variabilidade será necessário considerar suportes horários mais restritos que podem ir, eventualmente, até ao segundo, e sobre os quais se vão identificar as Regras de Associação, em especial se as previsões forem realizadas para o curto prazo.

Uma vez que a exposição global ao *media* apresenta irregularidade e o seu padrão de variabilidade não torna evidente uma forte influência dos conteúdos, as análises inicialmente devem ser consideradas em função das variáveis temporais. A maior ou menor necessidade de considerar as características dos conteúdos dependerá dos modelos explicativos da exposição que condicionem em maior grau o comportamento do painel.

II. Identificação do conjunto de programas a que os telespectadores se encontram expostos, independentemente da frequência de exposição e a identificação de segmentos

Uma vez que estes segmentos não apresentam hábitos de consumo televisivo, em cada segmento horário seleccionado, cada um dos telespectadores poderá ser caracterizado por várias sequências de exposição que podem variar significativamente no decorrer do tempo. Se considerarmos a totalidade da emissão diária, os telespectadores podem apresentar um número e uma constituição mais variável das suas sequências diárias no decorrer do tempo, mas, também, poderá não existir qualquer exposição. Será necessário extrair uma sequência por telespectador e, eventualmente, por segmento horário quando se recorre a intervalos de tempo com uma menor duração, por exemplo, com apenas alguns segundos. Devemos identificar uma sequência, que seja o mais representativa possível do comportamento mais característico de determinado telespectador, diferenciando o período semanal e de fim-de-semana e os períodos de férias em função da estação do ano.

O consumo televisivo apresenta um carácter sazonal que deve ser considerado nas análises. Atendendo aos resultados alcançados, estas análises devem ser conduzidas, preferencialmente, em função das faixas horárias. Pelo menos em dias laborais a alteração horária dos programas incute uma forte variabilidade aos valores dos *ratings* e os telespectadores não esperam pelos conteúdos transmitidos, quando estes são alterados. Sem dúvida, que prevalece a estrutura horária da programação. A exposição ao *media* corresponde à variável que condiciona a duplicação primária das audiências e é insuficiente para compreender a evolução dos *ratings*. Verificámos a existência de uma considerável correlação entre os valores de *ratings* e das taxas médias de audiências para o total da TV.

Telespectador 1 ... N	Sequências	
	Dia	Características
	Segunda	$\Delta t_{i1} + \Delta t_{i2} + \Delta t_{i3} + \dots$
	Terça	$\Delta t_{i7} + \Delta t_{i8} + \Delta t_{i9} + \dots$
	...	...
	Domingo	$\Delta t_{i21} + \Delta t_{i22} + \Delta t_{i23} + \dots$

As sequências mais características podem variar na sua dimensão, considerando um crescente número de faixas horárias.

Quadro VI.4.5.1.1 – Identificação dos suportes por telespectador e dia e construção das sequências mais representativas.

Fonte: Elaborado pela autora.

As Regras de Associação correspondem a proposições probabilísticas sobre a ocorrência de determinados estados nas bases de dados. As Regras mais relevantes, no âmbito do planeamento de *media*, serão as que apresentam uma maior cobertura do alvo ou da base de dados em análise (no sector, considera-se como satisfatório alcançar 50% do público-alvo) e os custos influenciam o número de suportes. Devem-se considerar as Regras com as “melhores” frequências condicionadas, acurácia ou grau de precisão para o horizonte temporal considerado e atendendo ao conjunto de Regras que foi viável extrair dos dados. Podemos proceder à identificação dos suportes a que os telespectadores se encontram mais expostos, independentemente da sua frequência em determinado período e, posteriormente, proceder à identificação de segmentos de telespectadores com base em algoritmos de *clustering* de sequências.

Cada telespectador apresenta várias sequências de exposição, que podem variar no decorrer do tempo. O conjunto de programas e de canais a que assistem é variável por segmento e por indivíduo. Devemos extrair uma sequência temporal por telespectador e por segmento horário, se aplicável, uma vez que podem ligar e desligar o TV várias vezes no decorrer do dia. Os anunciantes podem querer fazer incidir a campanha publicitária em determinados períodos que são função dos custos e do perfil das suas audiências. As Regras de Associação dizem respeito a uma sessão, enquanto que as Regras Sequenciais podem estabelecer relações entre sequências ou séries, ou seja, várias sessões num dia intercaladas por determinado intervalo de tempo ou relacionando sequências observadas em diferentes dias, ainda que considerando um curto espaço de tempo.

As Regras de Associação podem avaliar as frequências entre suportes horários que não sejam adjacentes e que correspondam a diversas sessões televisivas dos telespectadores no decorrer de um dia. As Regras podem apresentar diversos níveis de abstracção relacionando, por exemplo, as tipologias de conteúdos e podem incluir atributos numéricos ou trabalhar com variáveis que apresentem mais de dois valores. As Regras Sequenciais são mais relevantes no âmbito programático, por exemplo, identificando o número de telespectadores que assistem a determinado programa e que, também, assistem a um segundo programa ao fim de algumas horas ou alguns dias.

Estes algoritmos incidem em bases de dados onde se identificam as variáveis temporais associadas a cada sessão, não sendo neste caso necessário conduzir as análises por segmento horário. Procura as sequências, ou seja, um conjunto de suportes ordenados em função do tempo. A sua cobertura será o número de pessoas que contêm essa sequência, permitindo a quantificação de um dos indicadores que é essencial para o planeamento de *media*, na medida em que esses suportes permitem expor os anúncios a determinada percentagem do público-alvo. Seleccionam-se os suportes com determinada cobertura mínima e calculada face a cada telespectador e transformam-se as sequências originais deixando apenas os suportes frequentes.

Calculam-se as sequências de diferentes dimensões que permitem uma cobertura mínima. Pretende-se as sequências de dimensão mínima, mas que apresentem uma cobertura aceitável e com uma frequência condicionada que, também, permita um grau de duplicação de audiências que seja considerado aceitável.

Quando as análises são conduzidas em função dos suportes horários podem ser aplicadas as Regras de Associação que nos proporcionam uma visão estática dos padrões identificados. A identificação dos suportes horários permite-nos ter uma maior percepção sobre o tempo que separa as eventuais sessões televisivas que possam existir com uma maior frequência num dia. As sessões podem incidir em diferentes canais, e, logo, os suportes horários pedem uma identificação dos canais a que dizem respeito, por exemplo, $T_{20-30}SIC$ associado a $T_{40-50}TVI$.

Uma vez identificados os suportes horários mais frequentes, e atendendo às alterações que, eventualmente, se possam verificar nos conteúdos transmitidos ou nos horários de transmissão dos programas podemos, eventualmente, considerar as sequências programáticas mais características. Por exemplo, se viu $Prog\_x\_SIC$ nesse segmento horário poderá com uma probabilidade p_1 ver o $Prog\_y\_SIC$ ou $Prog\_w\_TVI$, também, nesse segmento horário já identificado no espaço de tempo de z dias, horas, minutos. Contudo, eventualmente, estas análises devem ser realizadas por segmentos identificados e, não, apenas para cada telespectador. As eventuais regularidades devem ser consideradas para um colectivo de indivíduos.

Com as variáveis simbólicas podemos avaliar a dispersão ao nível das tipologias e programas que os suportes horários apresentam, por exemplo, $T_{20}SIC$ caracteriza-se por [20% (Informação), ... , 15% (Divertimento)] ou [20% (Espaço Informativo), 5% (Futebol), 10% (Filme_ Y), ... , 15% (Telenovela_X)].

Uma vez que o anunciante vai adquirir tempo comercial associado a determinado suporte programático e verificamos que a associação entre os respectivos valores de *ratings* é elevada, quer linearmente, quer do ponto de vista não linear, as análises podem ser conduzidas em função dos suportes horários programáticos, que, assim, podem apresentar uma maior duração. Verificamos que se verificam transições entre os canais no decorrer dos programas, mas não corresponde a um dos factores mais relevantes para a quantificação da duplicação primária das audiências, ainda que este indicador seja por si só insuficiente para compreender a evolução dos *ratings*. Será necessário avaliar qual é a dimensão dos suportes horários mais adequada, ainda que, no sector, não seja frequente considerar valores inferiores aos 15 minutos quando se pretende a quantificação dos *ratings*. Também, será necessário avaliar a evolução das audiências no decorrer de um programa. Posteriormente, as análises realizadas no contexto comercial permitem-nos obter uma maior percepção do perfil do comportamento do segmento face aos espaços comerciais e condicionadas pelos factores de planeamento, como, por exemplo, duração do intervalo e número de inserções para a observação a prever.

Em função do grau de heterogeneidade dos telespectadores devemos realizar o *clustering* das Regras mais representativas. Podemos recorrer a algoritmos de densidade ou considerando variáveis que caracterizem a duração da exposição e o número de transições entre os canais e, posteriormente, ao *clustering* hierárquico e/ou ao algoritmo K Média. Para que as conclusões sejam válidas deverá existir a maior homogeneidade de comportamentos e devem ser retiradas observações atípicas das bases de dados antes que se calculem os modelos locais. Corresponde a uma etapa prévia que poderá contribuir para uma maior precisão dos resultados face aos resultados de considerar a totalidade da amostra.

Com os algoritmos de densidade de *clustering* de sequenciais podemos recorrer a medidas de semelhança entre as Regras que são função do grau da sua compactacidade e o factor de ponderação de cada suporte irá depender do valor da sua frequência nesse agrupamento.

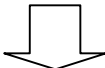
O algoritmo de *clustering* terá que permitir uma análise rápida e computacionalmente eficiente dos dados, minimizando a intervenção do utilizador. A transformação da base de dados original num quadro booleano que representa a ausência ou presença de determinado suporte na sessão viabiliza a aplicação de algoritmos tradicionais, mas a sua dimensão e a alta dimensionalidade, por vezes, torna este processo ineficiente.

As Regras de Associação permitem a identificação dos suportes que são seleccionados em conjunto com uma maior frequência e num mesmo intervalo de tempo, independentemente da ordem associada à exposição. As Regras Sequenciais utilizam-se quando as selecções não se verificam num mesmo instante, por exemplo, quando ocorrem em diferentes sessões no decorrer do dia.

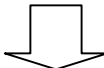
Uma vez que outros segmentos, para além dos enumerados anteriormente podem emergir, a sequência de abordagens metodológicas a ser adoptada deverá em primeiro lugar proceder à segmentação do painel e, posteriormente, para cada subamostra deve ser calculada a sequência de selecções programáticas ou de faixas horárias mais representativas ou por análise dos respectivos índices de audiências ou de duplicação entre as faixas horárias e entre os episódios de um programa.

Os Apêndices F.3. e F.4 abordam a problemática da duplicação das audiências e nos Apêndices F.1, F.2 e F.6 consta a análise dos índices de audiências programáticos, quer do ponto de vista univariado, quer multivariado. O seguinte *Quadro VI.4.5.1.2* resume a sequência de análises proposta. Conseguir identificar todos os conjuntos frequentes corresponde a um processo lento, e, logo, o algoritmo inicia com a identificação dos conjuntos compostos por apenas um item. Os conjuntos frequentes são compostos de elementos que, também, são frequentes. Posteriormente, constroem-se os conjuntos com dois elementos que incluem elementos frequentes e avaliam-se quais destes, também, são mais frequentes, e, assim, sucessivamente.

<i>Telesp.</i>	<i>Dia</i>	<i>Sequência/ Dia</i>
1	1	T20, T21, ..., T30 (Sessão única)
	2	T50, T51, ..., T70 (Sessão única)
	...	...
	n	T60SIC, T62SIC, T63SIC (Sessão1) T80TVI, .., T90TVI (Sessão 2)



<i>Ind.</i>	<i>Sequências</i>
1	T10, ... , T20
2	T25SIC, ... , T40TVI
...	...
n	T30, ... , T50



<i>Identificação de Segmentos</i>	<i>Sequências Para cada segmento</i>
Segmento 1	T2RTP1, ..., T10SIC
...	...
Segmento N	T30, ... , T50

Quadro VI.4.5.1.2 - Identificação de sequências de exposição.

Fonte: Elaborado pela autora.

III. Construção de uma Sequência Característica Por Telespectador

Ex1: $T_{20} \dots T_{90}$

(Totalidade das exposições)

Alguns suportes podem ser pouco frequentes no global, e, logo, deve ser restringida a sua dimensão

Ex2: $T_{50} \dots T_{80}$ ou $T_{30SIC} \dots T_{50TVI}$

Pode representar determinada percentagem do número total das exposições, por exemplo, 50%. Deverá existir um “trade-off” entre os custos que se enquadram no orçamento disponível e o número de suportes.

Convém conhecer as sequências mais vastas, uma vez que devem ser seleccionados os suportes com equivalentes graus de exposição, mas com menores custos. Alguns suportes, também, podem não estar disponíveis, por terem sido adquiridos por agências e anunciantes concorrentes.

Uma maior percepção da evolução das audiências comerciais permitirá uma selecção dos melhores tempos publicitários em determinado intervalo. Por conseguinte, devemos extrair uma dimensão mínima de suportes, mas que seja representativa e que maximize a frequência para valores considerados aceitáveis, assim como a confiança e o “Lift” de uma Regra de Associação.

IV. Análises de caracterização de cada segmento e da exposição global

Finaliza-se com as análises de caracterização de cada segmento e da exposição global, independentemente das manifestações mais concretas em determinada semana. Será necessária a identificação de Regras Sequenciais ou de Associação ou de Dependência mais representativas para cada segmento previamente identificado. Podemos recorrer ao *Link Mining* e ao MDS para visualizar as associações existentes e o grau de duplicação das audiências.

Com o *Link Mining* avaliamos que suportes são vistos em conjunto em maior grau e os suportes que não apresentam ligação, e que, portanto, não são complementares, assim como, também, recorrendo às análises MDS. Permite-nos identificar quais são os suportes que se associam com os suportes que se destacam por apresentar uma maior frequência de exposição conjunta com a restante amostra. Uma vez que no decorrer do tempo podem existir alterações na programação ou faixas horárias seleccionadas, a implementação de Regras Sequenciais em ambientes dinâmicos constitui uma via de actualização dos resultados e na identificação das alterações observadas nas Regras, uma vez que a sazonalidade e a alteração da grelha de programas influênciam os padrões de exposição. Podem ser aplicados algoritmos adaptados a meios dinâmicos, uma vez que as Regras identificadas podem variar no decorrer do tempo.

A principal fonte de variação deverá corresponder, essencialmente, à introdução de novos programas, e nesses contextos será necessário verificar quais são as associações que permanecem ou se extinguem e quais as novas que vão emergindo.

Também, podemos querer identificar como as associações diferem das Regras históricas, a frequência observada nas alterações que são observadas e que programas apresentam maior ou menor vulnerabilidade. Portanto, podem ser utilizados os algoritmos que permitem a agregação periódica de blocos de transacções, e analisam os dados através de janelas temporais, assim como algoritmos mais rápidos através da utilização de técnicas incrementais, procura em paralelo e na selecção das associações mais relevantes segundo determinadas heurísticas.

As alterações podem ser mais graduais quando derivam da sazonalidade do consumo televisivo, ou mais rápidas quando ocorrem consideráveis alterações na grelha, apesar de ser inegável a influência estrutural que torna a exposição global ao *media* num dos principais factores que conferem variabilidade aos dados e que, frequentemente, se apresenta como independente dos conteúdos. Verificando-se uma exposição ao *media*, poderá ser menor a variabilidade ao nível da selecção dos suportes, que por sua vez devem estar associados às sequências previamente identificadas. Uma maior influência das variáveis de gratificações e usos do *media*, como, por exemplo, o interesse por determinado conteúdo poderá contribuir para modificar os suportes horários mais habituais e, posteriormente, também as selecções.

Após a apresentação de um novo bloco de observações, os sistemas de *Data Mining* devem actualizar os valores das frequências e verificar se as variações são consideráveis, o que poderá ser realizado recorrendo a medidas de semelhança entre os dois conjuntos de suportes. A dimensão da janela que define a amostra sobre a qual incidir as análises será tanto menor quanto maior for a variabilidade e a dimensão de cada bloco poderá ser diferente no decorrer o tempo. Cada conjunto de indivíduos pode ser analisado segundo várias vertentes, por exemplo, quanto à caracterização e diversidade de canais, períodos, tipologias e programas seleccionados; grau de dispersão de resultados observado no segmento, ou seja, se os indivíduos apresentam uma grande semelhança ao nível da exposição televisiva e a evolução observada nas Regras identificadas.

Podemos recuperar os dados iniciais e verificar se cada segmento abrange indivíduos que apresentam uma grande variabilidade nas respectivas Regras ou não. Devemos identificar comportamentos de lealdade programática, tipológica e ao canal; tendências para se observarem os efeitos de herança, a repetição de selecções e as exposições repetidas; o grau de sensibilidade face à introdução de novos conteúdos na grelha; o grau de complementaridade entre os conteúdos seleccionados; a caracterização da composição sociodemográfica dos telespectadores, entre outros aspectos. Esta perspectiva de análise permite contribuir para a previsão das audiências, mas, também, para a caracterização da exposição televisiva dos segmentos que apresentam uma maior irregularidade no consumo televisivo.

A partir destes dados, podemos avaliar quais são os programas que lhe estão associados, considerando, por exemplo, variáveis simbólicas onde consta para o período em análise que percentagem de observações corresponde a diferentes programas ou tipologias programáticas. A extracção de Regras sequenciais podem ser alcançada através da seguinte sequência de análises representada nos *Quadros VI.4.5.1.3-5*.

Id_Telespectador Ou Segmento	Id_Tempo	Itens
1	1	T20
1	2	T50
2	3	T10, T20
2	4	T40, T42, T43
2	2	T20
3	1	T30, T40
3	3	T50, T20

Inicialmente devemos identificar as sequências por individuo e tempo e calcular as sequências para cada telespectador (*Quadro VI.4.5.1.3*).

Quadro VI.4.5.1.3 – Id. suportes por telespectador.
Fonte: Elaborado pela autora.

Aos conjuntos de suportes temporais com um valor de cobertura considerado satisfatório atribuir uma identificação (*Quadro VI.4.5.1.5*) e proceder ao cálculo das sequências transformadas (*Quadro VI.4.5.1.4*).

Id_Seg.	Itens	Sequencia Transformada
1	(T20), (T50), (T30, T40)	1,4,5
2	(T10, T20), (T40, T42, T43), (T20)	1,3
3	(T30, T40), (T50, T20)	5,1,2,3,4

Quadro VI.4.5.1.4 – A construção de sequências transformadas. Fonte: Elaborado pela autora.

Itens com cobertura mínima de 2	Id_Sequência
(T20)	1
(T30)	2
(T40)	3
(T50)	4
(T30, T40)	5

Quadro VI.4.5.1.5– A identificação de sequências com uma cobertura mínima.
Fonte: Elaborado pela autora.

Segue-se a identificação das sequências que apresentam uma cobertura mínima e a sua redução, eliminando as subsequências existentes a partir das sequências que apresentam um maior número de elementos. Também, podem ser consideradas restrições temporais ao nível da duração das sequências, entre outras.

V. A selecção de suportes numa óptica de planos de *media*

Com base nos conteúdos programáticos identificados no item anterior, podemos construir o plano de *media* de referência e do qual resultará o plano final por aplicação de algoritmos de optimização. Numa óptica de construção de planos de *media* devemos seleccionar os suportes que constam das sequências mais características do público-objectivo pretendido ou com base em análises aos indicadores sociodemográficos que constam das bases de audimetria e incidir as análises de previsão nesses suportes. A distribuição de contactos será, assim, a que for “característica” de determinado segmento de telespectadores, mas, posteriormente, poderá ser melhorada atendendo aos objectivos da campanha publicitária, recorrendo aos algoritmos de optimização.

A selecção dos suportes, também, se deve basear na análise dos índices de adesão, afinidade, atractividade e fidelidade. Do ponto de vista metodológico podemos recorrer às análises que constam do *Apêndice F.5*.

As análises MDS permitem que as associações sejam realizadas recorrendo à Regressão linear ou alternativamente à análise Discriminante linear e à análise de *Custers*. A análise de Correspondências permite a identificação dos suportes que se associam em maior grau com determinados segmentos e que devem ser seleccionados para integrar os planos de *media*.

Quando os programas apresentam um maior número de transmissões devemos recorrer à Análise Simbólica que te em atenção a variabilidade dos dados. Interessa seleccionar os suportes com maiores níveis de exposição, mas que simultaneamente apresentem uma menor variabilidade. Devemos seleccionar as tipologias e os programas que apresentem maiores valores de adesão e de fidelidade por parte do seu público. A identificação do perfil de audiências de cada canal contribui para uma maior compreensão da evolução dos principais indicadores, atendendo aos segmentos que foram identificados como apresentando evoluções mais diferenciáveis face ao universo televisivo e como sendo mais variáveis.

VI.4.5.2. A análise da exposição aos suportes seleccionados e as teorias explicativas da exposição televisiva

Na segunda etapa das análises proceder à análise de previsões mais concretas para determinada semana, ainda que nas Ciências Sociais as previsões são classificadas como inexactas, onde apenas é possível avaliar direcções, tendências, propensões e probabilidades de alterações de estado. As análises anteriores permitem avaliar quais são os suportes mais frequentes. Porém, em determinada semana ou dia poderá não se verificar a exposição aos mesmos, uma vez que não existe uma considerável regularidade do consumo televisivo. A exposição pode variar no decorrer do tempo face aos conteúdos seleccionados e respectiva frequência que dificilmente será previsível. Verificamos que a exposição ao *media* será a principal variável responsável pela irregularidade do consumo televisivo e, por outro lado, a vontade e a disponibilidade são muito difíceis de prever.

Por conseguinte, será necessário tentar construir algumas Regras de previsão, classificação e de avaliação de dependência entre as variáveis explicativas com um maior potencial de previsão. As variáveis podem constar das bases de audimetria ou ser calculadas a partir destes dados, assim como podem ser calculadas a partir de inquéritos que permitam avaliar a envolvente social, os factores individuais e estruturais que determinam a exposição ao *media* e as suas interdependências.

Uma possibilidade será recorrer a Redes de Bayesianas (*Jensen, 2001; Han & Kamber, 2001*) e a Regras ou a Árvores de *Data Mining* Multirrelacional (*Kramer & Widmer, 2001*) para avaliar a inter-relação observada entre os diferentes agentes e os comportamentos de selecção. A construção de uma base de dados relacional considerando apenas a informação que consta das bases de audimetria é insuficiente, e, também, não justificaria recorrer a abordagens relacionais. Uma vez identificados para cada segmento, o conjunto de programas que poderá vir a seleccionar em determinada semana, independentemente da frequência e da diversidade de conteúdos, será necessário aprofundar os factores que condicionam a sua exposição ao *media* e, assim, tentar prever manifestações mais concretas.

Os modelos explicativos da exposição são essenciais neste contexto de análise e na presente tese avaliaram-se alguns dos factores estruturais que influenciam o consumo televisivo, uma vez que segundo a bibliografia da especialidade, a sua influência supera a que resultaria das variáveis individuais, inclusive em mercados caracterizados por uma grande diversidade de canais disponíveis (*Cooper & Tang, 2009; Wonneberger et al., 2011; Wonneberger et al., 2012*).

De acordo com *Wonneberger et al. (2012)*, por análise de duas décadas de dados de audimetria Holandeses que caracterizam a exposição apenas aos espaços informativos, verifica-se uma crescente exposição a esta tipologia de programas que se encontra maioritariamente influenciada pelos factores estruturais e de contexto, mas em muito menor grau por parte das motivações dos telespectadores.

Neste trabalho, os autores demonstram que apesar da maior oferta televisiva informativa, coexistindo o sector público e o privado, estes programas conseguem apresentar uma elevada cobertura das audiências e que a especialização é reduzida, ou seja, a tendência para incidir num número restrito de conteúdos. A exposição poderia resultar de interesses e de uma motivação por ver os conteúdos informativos, mas, também, de uma selecção não intencional influenciada pelas estratégias de programação e que determina a exposição aos conteúdos que não correspondem aos programas e tipologias realmente pretendidos.

Os autores avaliam a influência de alguns factores motivacionais, como, por exemplo, o interesse político, a taxa de transição entre os canais e a prontidão da selecção. Como factores de contexto consideram a disponibilidade, o repertório de canais, os efeitos de “Lead_In”, “Lead_Out”, o “co-viewing” e algumas variáveis sociodemográficas. As principais conclusões que são referidas correspondem a que um crescente número de canais não levou a um decréscimo na exposição global aos espaços informativos, apesar da fragmentação das audiências e que determinados segmentos diferem significativamente ao nível da evolução da intensidade do seu consumo televisivo. Considerando análises de Regressão Binomiais Negativas concluem que a influência estrutural prevalece e que as variáveis que mais destacam são a disponibilidade e a prontidão da selecção, assim como a idade.

O factor “Lead_In” apresenta a maior evolução ao nível da sua crescente repercussão no consumo, enquanto o repertório de canais, o “co-viewing”, o sexo e o nível de instrução apresentam uma influência moderada. Os telespectadores assistem aos espaços informativos simplesmente porque assistem à TV. Quanto maior for o número de alternativas, verifica-se uma menor tendência para evitar os conteúdos informativos por transição entre os canais. As pessoas reagem a uma crescente complexidade da oferta televisiva, fazendo incidir as suas seleções num número restrito de canais ou num canal, ou seja, observa-se uma nítida tendência para o desenvolvimento de comportamentos de lealdade ao canal. Destaca-se o sector público, assim como se verifica a nível nacional, uma vez que lhe é atribuída uma reputação de maior fiabilidade e isenção.

As audiências dos espaços informativos resultam de uma combinação de factores ambientais e de uma audiência que essencialmente vê TV e que apresenta hábitos de consumo televisivo. Alguns destes resultados contradizem inúmeros trabalhos realizados em universidades norte-americanas.

Em *Wonneberger et al. (2011)*, por análise da exposição aos espaços informativos recorrendo à audimetria, os autores pretendem uma vez mais avaliar as influências estruturais e das variáveis de motivação. Como factores estruturais consideram a disponibilidade quantificada pelo número de horas de exposição ao TV, o efeito “Lead_In” e “Lead_Out” e o “Co-Viewing”. Nas variáveis de motivação incluem o interesse político, a preferência, a antecedência com que os telespectadores ligam a TV antes do início do espaço informativo e a taxa de transição entre os programas. Os autores concluem, uma vez mais, que os factores de contexto prevalecem largamente sobre os factores motivacionais, representando um efeito sete vezes superior. “However, situational influences should not be misunderstood as deterministic, rendering passive viewers. Viewers may deliberately relinquish their choice opportunities and leave their news exposure to chance.” (*Wonneberger et al., 2011, p.338*). A exposição será superior quanto maior for a disponibilidade, se assistirem em grupo, se assistirem ao programa anterior e sucessor no mesmo canal, se demonstrarem interesse e preferência por esta tipologia de programas, se transitarem, frequentemente, entre canais e se ligarem a TV num instante próximo ao seu início.

A exposição e a sua frequência depende obviamente do tempo total de exposição ao televisor, independentemente dos conteúdos que forem seleccionados. Os factores situacionais podem provocar grandes diferenças na frequência e no grau de permanência dos telespectadores no decorrer dos espaços informativos. Quem vê mais TV, também, assiste a um maior número de espaços informativos, num maior número de dias. Assistir em grupo a determinado programa e aos programas de “Lead_In” apresenta uma maior repercussão do que os factores motivacionais.

No conjunto de factores motivacionais considerado nas análises, os autores salientam a prontidão da selecção e referem a reduzida relevância dos interesses e que as preferências levam a um acréscimo de apenas 7 minutos na exposição face ao segmento que não demonstra preferência por esta tipologia de programas. Os efeitos dos factores motivacionais considerados não apresentam inter-relações.

O factor estrutural que determina em maior grau a exposição, ao nível da frequência de selecção, corresponde à disponibilidade, e ao nível do tempo de permanência das audiências destaca-se a variável que representa se assistem individualmente ou em grupo. Salientam que as interacções observadas entre as variáveis individuais e estruturais não viabilizam grandes melhorias na capacidade explicativa dos modelos desenvolvidos. Apenas identificam três interacções como sendo mais significativas. Os telespectadores que manifestam um menor interesse em política, assistem em maior grau aos espaços informativos se pretenderem ver o seu programa sucessor e se assistirem à TV em grupo. O interesse influencia a disponibilidade. Quem vê mais TV assiste aos espaços informativos durante mais tempo se existir interesse em temáticas da esfera política. A preferência contribui para uma maior influência da disponibilidade e do “co-viewing” e o interesse modifica a repercussão do efeito do programa sucessor.

Os resultados apenas confirmam parcialmente as interacções entre os dois subconjuntos de variáveis. As variáveis sociodemográficas apresentam uma maior influência do que as variáveis motivacionais. Relativamente à influência do grupo e do agregado familiar na exposição, em *Mora, Jason e Krider (2011)*, com dados de audimetria mexicanos constam algumas análises de Regressão. Os autores concluem que a tipologia, as variáveis psicográficas, o “co-viewing” dos programas anteriores são variáveis que condicionam a exposição em grupo perante determinado programa.

Assistir em grupo contribui para uma maior duração da permanência dos telespectadores no decorrer dos programas e para uma menor transição entre os canais. Salientam o cariz social da “experiência televisiva”.

No campo da duplicação das audiências, a nível agregado, e em especial da lealdade aos canais, em *Yuan e Ksiazek (2011)*, os autores que recorrem a dados de audimetria chineses referem que as audiências independentemente do número de canais existentes tendem a seleccionar os canais que apresentam maiores valores de *shares* de audiências e um maior número de conteúdos com um cariz dramático, uma vez que a TV será, essencialmente, um veículo de entretenimento no mercado Chinês. As estações televisivas, também, incidem numa concorrência por mimetismo. Verifica-se a existência de duplicação das audiências entre os canais e alguma tendência para se consolidarem repertórios de canais que coincidem, essencialmente, com os mais populares.

Verifica-se, também, uma concentração do mercado televisivo e uma inter-relação recíproca entre as estruturas e a actividade da audiência. As audiências apresentam diferentes padrões de consumo televisivo influenciadas pela envolvente e com o tempo os comportamentos colectivos podem alterar as estruturas em que os agentes operam. Também, salientam que apesar das audiências agregadas resultarem do somatório de telespectadores individuais, em função do nível de agregação dos dados os padrões observados diferem.

De acordo com *Sewell (1992)*, a agregação pode influenciar a estrutura e nesta capacidade reside a actividade da audiência. Apenas os dados agregados apresentam o potencial de influenciar a estrutura do sector, logo, as instituições monitorizam e respondem aos padrões do comportamento da audiência. As audiências são percebidas como agentes que viabilizam e se encontram restringidos pelas características estruturais do meio televisivo. Consequentemente, os modelos explicativos e de previsão devem ser desenvolvidos do ponto de vista agregado, ou seja, incidir nos fluxos de audiências sendo esta a perspectiva adoptada nesta tese. Devem-se considerar nos modelos, as variáveis sociodemográficas e psicográficas, os factores estruturais e alguns motivacionais, apesar de prevalecer a influência dos factores de contexto em detrimento dos motivacionais, segundo a bibliografia mais recente, assim como das respectivas interacções.

A repercussão de cada uma das variáveis certamente que depende do segmento em análise, tipologia de programa, faixa horária e canal. Dificilmente as Regras de previsão que possam ser identificadas vão coincidir para a totalidade dos suportes disponíveis e que podem integrar um plano de *media* que tem como objectivo divulgar com a maior eficácia as mensagens publicitárias.

Alguns factores estruturais e individuais a serem avaliados corresponderá, por exemplo, aos que constam do *Quadro VI.4.5.2.1.*, não sendo a sua enumeração exaustiva. Segundo *Wonneberger et al. (2012, p.69)*: “Future research should address the relative strength of motivacional and situational factors in increasingly complex choice situations.” Também, será necessário avaliar as interdependências existentes e em que contextos prevalece a influência de determinadas variáveis.

(1) A composição do agregado familiar e o grau de influência dos diversos agentes nas decisões de acesso ao *media* e nas selecções dos conteúdos. Avaliar quais são os telespectador que apresentam o maior poder de decisão e quais são os suportes que vão prevalecer em função dos contextos;

(2) Quais as características sócio-económicas-culturais, demográficas e psicográficas do indivíduo e do agregado familiar em que este se encontra inserido;

(3) Verificar se o telespectador apresenta um horário laboral e para quando vê TV e se desenvolve hábitos de consumo televisivo ou de disponibilidade para com o *media*, independentemente das selecções realizadas;

(4) Identificar se o telespectador têm interesse em determinado canal, programa, tipologia. Avaliar se a selecção está em maior grau associada a motivações instrumentais ou ritualísticas e qual a sua estabilidade no decorrer do tempo ou face às características da programação;

(5) Avaliar quais são as opções de *media* e de lazer que podem constituir uma alternativa concorrente e a repercussão na selecção quando está disponível uma maior diversidade de canais. Com a televisão digital o número de opções poderá no futuro aumentar, o que poderá contribuir para uma maior variabilidade dos indicadores de audimetria e da exposição televisiva;

(6) Verificar se o telespectador apresenta o hábito de ver determinado programa e se também assiste com regularidade ao programa anterior e/ou sucessor. Avaliar o grau de duplicação das audiências entre os suportes publicitários.

Factores motivacionais e psicológicos	Variáveis sociodemográficas entre outras	Factores estruturais
Interesses	Sexo	Disponibilidade
Preferências	Faixa etária	Repertório de canais
Motivações para aceder ao <i>media</i>	Região geográfica	“Lead_In”
Taxa de transição	Classe social	“Lead_Out”
“Appointment viewing”	Grupo ocupacional	“Co-viewing”
Conhecimento da grelha	Nível de instrução	Lealdade tipológica
	Existência de TV_Cabo	Lealdade ao canal
	Responsável pelas compras do lar	Lealdade programática
	Taxa de consumo de jornais	Exposições Repetidas
	Religião	Acesso ao <i>Media</i>
	Nacionalidade	
	Actividades de lazer	

Quadro VI.4.5.2.1 – Algumas variáveis que determinam a exposição televisiva e a selecção dos conteúdos. Fonte: Elaborado pela autora.

A inferência indutiva poderá viabilizar a previsão de observações futuras a partir dos dados históricos se existirem regularidades que podem ser identificadas recorrendo a algoritmos Bayesianos e Multirrelacionais.

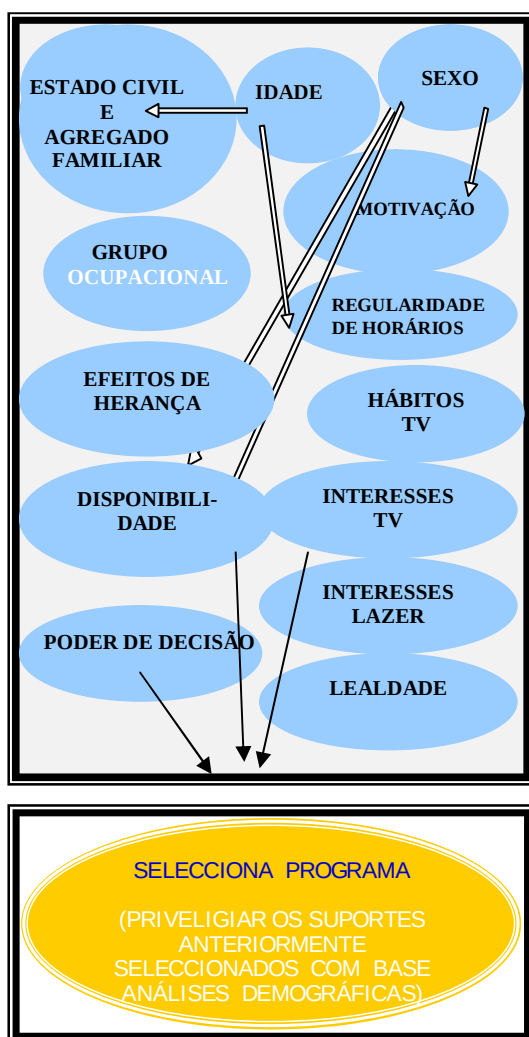


Fig. VI.4.5.2.1 - Rede de Bayes para a identificação de associações entre as variáveis que determinam a exposição televisiva.

Fonte: Elaborado pela autora.

Os factores que podem condicionar a exposição são inúmeros, e, logo, previamente será necessário identificar como a sua associação pode determinar a selecção de uma tipologia de suportes ou faixas horárias. A aplicação das Redes Bayesianas (Jensen, 2001; Han & Kamber, 2001) permite a avaliação dos factores que podem condicionar a selecção dos suportes considerados como mais representativos. Apesar dos pressupostos mais simplificados do *Naive Bayes*, o método é considerado como sendo competitivo face aos métodos alternativos e permite avaliar as dependências probabilísticas e estruturais entre as variáveis (Fig. VI.4.5.2.1). As Redes lidam com incertezas e correspondem a modelos descritivos e predictivos, nos quais recorrendo à teoria da probabilidade quantificam as incertezas.

Estas metodologias identificam relações de independência ou relevância entre as variáveis e representa de forma explícita as relações existentes. Realizam inferências a partir dos dados e viabilizam o cálculo de probabilidades associadas a cada uma das hipóteses possíveis. Apenas as Ciências Naturais estão facilmente apoiadas na experimentação e na observação. As Redes permitem avaliar, a partir de determinados dados de entrada, a probabilidade de selecção em função de determinadas probabilidades definidas *a priori*.

A modificação da probabilidade de uma variável altera a distribuição das restantes variáveis. Permite a selecção das variáveis mais relevantes e, posteriormente, podemos proceder à construção de uma base relacional, mas que transcenda os indicadores e a informação que consta das bases de audimetria. Uma base de dados que apenas considere os indicadores e as variáveis de audimetria seria, a título ilustrativo, a que se descreve na seguinte *Figura VI.4.5.2.2*. Verificamos que existe a necessidade de ser construída uma base de dados mais abrangente, ao nível dos factores estruturais que determinam o consumo televisivo. Se a tipologia de programa em análise apresentar um público mais activo, também devemos incluir variáveis enquadradas na perspectiva de gratificações e usos do *media* e a respeito da sua interacção com a vertente estrutural. Esta informação terá que resultar de um inquérito a ser realizado ao painel.

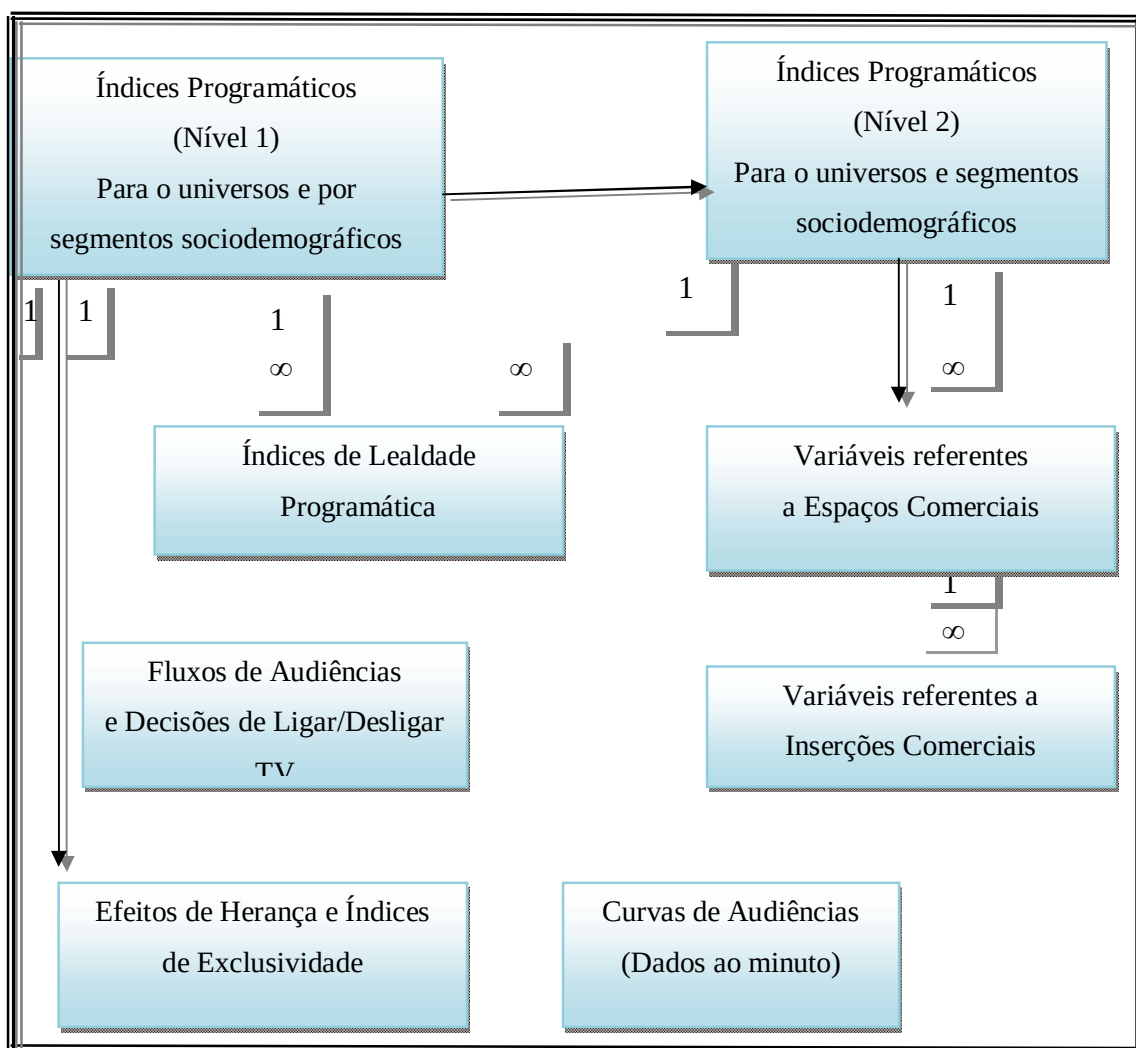


Fig. VI.4.5.2.2 – Base de dados relacional hipotética considerando a informação que consta das bases de audimetria. Fonte: Elaborado pela autora.

Podem ser calculadas inúmeras variáveis estruturais a partir das bases de audimetria, após a análise das sequências que caracterizam os telespectadores que integram o painel. Por exemplo, com a audimetria podemos saber qual será a constituição do grupo que rodeia um telespectador. Do ponto de vista do *Data Mining* Multirrelacional existem algoritmos de Árvores de Classificação (Kramer & Widmer, 2001; Atramentov, 2003; G-Feng, Jing & Wei-Feng, 2007) e Regras de Associação (Valêncio, Oyama, Ichiba & Souza, 2011), entre outras abordagens. Segundo a bibliografia da especialidade, em análises de Regressão onde é necessário considerar, essencialmente, atributos numéricos o desempenho destes algoritmos fica muito aquém dos resultados que decorrem dos algoritmos proposicionais (Knobbe, Siebes & Van der Wallen, 1999). Contudo, podemos tentar recorrer a Árvores de Classificação para tentar promover a previsão. Uma vez que esta tipologia de algoritmos produz melhores resultados quando as variáveis são do tipo nominal podemos considerar como variável dependente se existe selecção de determinado programa, tipologia ou canal em determinada faixa horária.

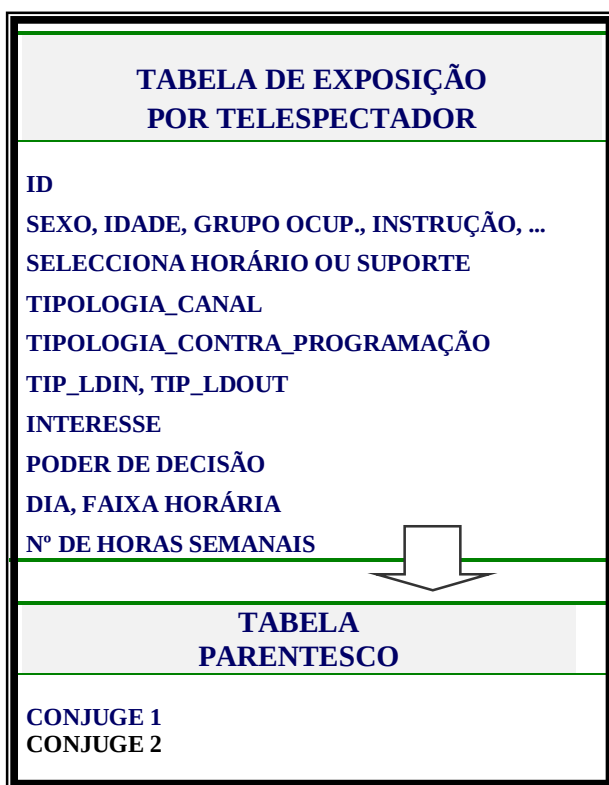


Fig. VI.4.5.2.3 – Representação de base de dados relacional para a identificação de Regras multirrelacionais. Fonte: Elaborado pela autora.

O DM Multirrelational (MRDM) (Dzeroski & Lavrac, 2001; Dzeroski, 2005) recebe contributo do *Inductive Logic Programming* (ILP) (Nienhuys-Cheng & Wolf, 1997), da Teoria das bases de dados, *Knowledge Discovery*, podendo-se utilizar directamente a representação relacional. Baseia-se em *queries* de *Structured Query Language* (SQL), em vez de na Programação Lógica. Os padrões relacionais são expressos em subconjuntos de lógica relacional, que inclui predicados, por exemplo, o parentesco e variáveis (C1, C2) que não estão presentes na lógica proposicional.

Os padrões relacionais acabam por ser mais expressivos do que os algoritmos proposicionais. Podemos proceder à derivação de Regras Multirrelacionais do tipo: EXPOSIÇÃO (C1, IDADE1, ...) PARENTESCO (C1, C2) $\wedge$ EXPOSIÇÃO (C2, IDADE2, ..., SEL_S) $\wedge$ INTERESSE=TELENOVELA $\wedge$ PODER_DEC=S. Essa relação descreve se alguém apresenta uma maior exposição a uma faixa horária ou programa se estiver casado com alguém que se interessa por telenovelas e com maior poder de decisão sobre as selecções a realizar em determinado momento.

Segundo *Almeida, Madureira e Cruzeiro (1973)*, a explicação causal nas Ciências Sociais traduz-se no enunciado de múltiplas condições determinantes cuja verificação simultânea acaba por conferir um determinado grau de probabilidade à ocorrência de determinados resultados. Às Regras podem ser associadas determinados valor de probabilidade.

Na realidade o conhecimento nas Ciências Sociais não se deixa apropriar segundo o modelo da causa explicativa isolada, pontual e susceptível de representação precisa através de uma função matemática. Apenas probabilidades podem ser calculadas, as quais podem diferir significativamente, e portanto podem estar justificadas nas áreas em que será relevante descartar as situações menos prováveis de ocorrer. De acordo com *Bunge (1981)*, as Ciências Sociais beneficiam com um estudo teórico e empírico mais aprofundado das probabilidades, permitindo, assim, prever a direcção da evolução mais provável dos acontecimentos. Segundo o autor, isto é, o que se faz em física atómica e na vida quotidiana, com a particularidade que neste último caso a projecção estocástica é geralmente qualitativa e não quantitativa e sem aprofundar ou debilmente fundamentada e sem serem baseadas em teorias científicas.

Verificamos que existe um potencial de aplicação de metodologias do âmbito do *Data Mining* Multirrelacional e da Estatística Relacional quando se pretende a análise das bases de dados de audimetria.

A sua principal mais-valia em relação aos algoritmos proposicionais consiste em considerar as relações entre as diversas tabelas e avaliar em que medida as inter-relações observadas permitem compreender a evolução da variável critério e realizar previsões desde que associando probabilidades a determinados cenários.

VI.4.6. A construção do plano final de *Media*

Após o desenvolvimento dos modelos de exposição será necessário recorrer à optimização, para ter em atenção o orçamento disponível e que condicionará fortemente a solução final. Actualmente, existe um grande descontentamento no sector face aos modelos existentes, sendo essencial o *Know-How* e a experiência dos planificadores para o seu desenvolvimento (*Makienko, 2012*):

Later, *media* planning industry experienced a significant growth in the use of reach and frequency optimization models that tried to substitute *media* planners. However, computerized models did not help to resolve the frequency paradox. First, optimizations is mainly driven by cost per exposure and does not take into account lots of qualitative factors. In addition, all inputs include effective frequency levels should be determined by people. Recently, Cheong, De Gregorio and Kim examined national *media* planning practice and concluded that most of *media* planners are less satisfied with such models. (p.60)

Consequentemente, é necessário a identificação de perspectivas alternativas de análise. Nas etapas anteriores foi possível identificar os suportes que apresentam uma maior cobertura do público-alvo e com frequências estruturais e características dos seus comportamentos, assim como a identificação dos cenários mais prováveis de ocorrer em determinada semana e que permitem restringir a dimensão da base de dados, na qual incidir os algoritmos que permitem seleccionar o melhor conjunto de suportes para determinado anunciante e campanha.

A sequência de análises proposta, na presente tese, permite a identificação das sequências mais características e adequadas ao segmento-alvo e que podem ser avaliadas em função da sua cobertura e do grau de duplicação das audiências que apresentam.

Posteriormente será, essencialmente, necessário ter em atenção os custos que lhe estão associados, uma vez que já foram identificados os suportes horários, programáticos e comerciais mais adequados ao alvo da campanha em função da informação disponível e das previsões que podem ser realizadas nas Ciências Sociais.

Em vez de utilizar a base de dados de exposições que caracteriza o comportamento do mês anterior devem ser utilizados os suportes que foram seleccionados nas etapas prévias das análises e considerar os que apresentam uma maior probabilidade de exposição para as semanas em que incide a campanha publicitária. As probabilidades de exposição podem ser calculadas através de inúmeras perspectivas de análise que são função das características do segmento em análise e do grau de variabilidade dos indicadores e da exposição televisiva.

Uma vez que em geral o orçamento disponível é reduzido, sugerimos que devem ser seleccionadas as sequências de menor dimensão e com menores custos associados, mas que permitam uma razoável cobertura do mercado alvo e se necessário que no passado tenham apresentado os valores da distribuição de frequências considerado como mais adequado, ainda que essa informação não traduza o que se poderá verificar no futuro. No mínimo terá que ser avaliado o seu grau de duplicação das audiências.

Quando os valores de *ratings* para os suportes identificados apresentam uma maior estabilidade de valores podemos recorrer à previsão deste indicador e, posteriormente, calcular os níveis de GRP das campanhas publicitárias que condicionam o custo global de aquisição de determinado plano de investimento em publicidade televisiva. O valor do GRP de um plano resulta do somatório dos *ratings* de cada inserção comercial que o integra. Verificamos a existência de uma elevada correlação linear entre os valores de *ratings* comerciais e programáticos. As audiências acumuladas permitem a comparação da eficácia de vários planos de *media* alternativos.

As características da campanha será o que determinará a importância da cobertura versus frequência e o valor do OTS mais adequado. Segundo *Makienko (2012, p.61)*, cada marca deve encontrar o seu valor de frequência, sendo em geral este o aspecto que é privilegiado, uma vez que os anunciantes em geral dispõem de reduzidos orçamentos destinados ao investimento publicitário: “Media plans with low budgets seeking maximum reach won’t be able to overcome the threshold frequency level for the advertising to become effective.”

De acordo com o autor, na prática são definidos objectivos de alcançar uma frequência média de 5 a 6 contactos, sendo o valor mínimo observado de 3 e o máximo de 12, contudo quando a exposição é variável o seu valor não é previsível. Concluí referindo que um *software* não substitui a experiência e conhecimento dos planificadores.

Por conseguinte, após a selecção das sequências de suportes programáticos e/ou temporais mais representativas, através do *Link Mining* podemos avaliar quais são os eventos que apresentam maiores níveis de audiência duplicada e não duplicada. Podemos saber para cada conjunto de suportes a percentagem de audiência partilhada no passado.

Se as sequências apresentarem uma cobertura equivalente, mas, não, ao nível da duplicação das audiências, a selecção da melhor opção será uma função dos objectivos de Marketing e dos custos envolvidos. Uma maior audiência duplicada, ou seja, se existir um maior número de telespectadores ou um maior número de registos em que os telespectadores estão expostos aos vários eventos, denota que esses instantes ou conteúdos são de certo modo complementares, independentemente dos factores que determinaram essa mesma exposição. Assim, a sua influência conjunta será superior ao nível da divulgação da mensagem.

Quando a duplicação é muito reduzida poderá resultar de uma exposição pontual, e, logo, será mais ou menos interessante em função das características dos segmentos e do produto.

Alguns autores consideram que a publicidade é suficientemente persuasiva e que uma exposição será suficiente se os consumidores estiverem dispostos a adquirir esse produto, sendo neste contexto mais interessante optar por suportes que permitam alcançar um maior número de telespectadores. Neste contexto seleccionar-se-iam os suportes com baixos níveis de audiência duplicada. Porém, uma vez que os níveis de atenção influenciam fortemente a capacidade de incentivar a aquisição dos produtos (*Cannon & Riordan, 1994*), a duplicação das audiências não deverá ser excessivamente reduzida.

Por outro lado, alguns autores consideram que a publicidade tem a função de recordar e que a compra poderá ser realizada num momento posterior e que neste caso será mais relevante uma maior duplicação e frequência de exposição. Uma sequencia que apresente um maior número de pares de suportes com uma maior duplicação das audiências permite-nos identificar os que permitem alcançar uma maior frequência estrutural de exposição, ainda que sem quantificar o seu valor, mas que poderá ser avaliado para os dados históricos, como sucede nos modelos existentes.

O mais relevante será saber se contactam muito ou pouco com esses suportes e não um valor em concreto, seja este um valor médio ou uma distribuição de frequências de contacto. Na bibliografia da especialidade não está demonstrado a existência de um valor óptimo, mas, sim, de eventualmente vários valores que dependem de inúmeros aspectos qualitativos relativos à situação em análise, frequentemente, subjectivos, sendo a intuição e o *Know-How* do planificador determinante neste processo. A real repercussão de um valor de frequência nunca será correctamente quantificável devido aos níveis de atenção que lhe são direccionados, mas não só, uma vez que os telespectadores podem inclusive estar enfrente ao TV, mas realizar outras actividades em simultâneo. No limite, uma maior probabilidade de exposição deverá conduzir a melhores resultados, face a uma exposição menos provável por ter sido mais pontual no passado.

As campanhas podem ser construídas de diversas formas, entre as quais: o investimento contínuo, intermitente contínuo, intermitente crescente ou intermitente decrescente. O planificador deverá definir a duração dos períodos, a sua frequência e o valor do respectivo investimento. Em função das características da campanha, as sequências podem ser identificadas para diferentes horizontes temporais, mais ou menos alargados. Quanto maior for o período mais fácil poderá ser a identificação de um padrão de exposição para os segmentos que apresentam uma exposição mais variável, mas o seu valor será função dos objectivos de Marketing. Nomeadamente se pretendemos alcançar o maior número de telespectadores num curto espaço de tempo ou se a mensagem deverá ser continuamente repetida no decorrer de um maior horizonte temporal.

Avaliação de suportes		Avaliação de sequências de suportes
1ª SEMANA	Hierarquização dos suportes em função dos valores de duplicação e cobertura	EX. SEQ1:(Sessão_1) T30SIC - T40SIC, T41TVI - T50TVI, T52RTP
...	T30-t31: P(31/30), P(30), P(31) T89-T90: P(90/89), P(89), P(90)	SEQ(2):(Sessão_2) T70SIC - T80SIC, T81TVI - T90TVI
Nª SEMANA	NOTA: P(T _j /T _i) com i e j adjacentes ou não e podendo pertencer a diferentes sessões televisivas.	NOTA: Após a identificação das sequências devemos seleccionar os suportes que resultem das análises factoriais ao perfil das audiências e da fidelidade, lealdade aos conteúdos.

Quadro VI.4.6.1 – Identificação dos suportes a incluir nas grelhas de exposição onde incidem os algoritmos de optimização.

Fonte: Elaborado pela autora.

Podem ser extraídas Regras em que se define um período de horas ou dias a intercalar as selecções mais prováveis. Quando estão em análise segmentos que apresentam uma exposição mais irregular poderá ser identificado um padrão de exposição para um maior período e posteriormente redimensionado em função da sua probabilidade de ocorrência de semana para semana.

A informação que consta do quadro anterior permitirá a construção de uma grelha de exposição, sobre a qual irá incidir um algoritmo de optimização (*Craig & Ghosh, 1993*). Podemos hierarquizar cada par de suportes em função da sua cobertura e da duplicação das audiências, o que nos permitirá seleccionar as “melhores” sequências de exposição, complementando as análises com a análise do perfil das suas audiências. Para cada suporte devemos conhecer o seu custo de aquisição, cobertura do alvo, duplicação face aos restantes suportes e o perfil das suas audiências, podendo ser atribuído um maior factor de ponderação aos telespectadores ou lares que apresentem determinadas características quando se recorre aos algoritmos de optimização.

VI.5. OS MODELOS DE EXPOSIÇÃO PARA PROGRAMAS QUE GERAM OBSERVAÇÕES “ATÍPICAS”

Quando se transmitem programas que geram observações atípicas e uma maior interacção entre as séries temporais multivariadas referentes aos diversos canais será necessário considerar perspectivas metodológicas alternativas.

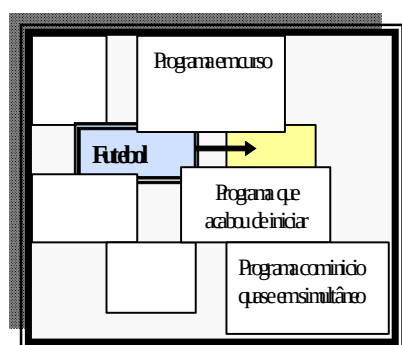
De acordo com *Nikolopoulos (2010)*, a repercussão dos eventos especiais na evolução das séries temporais deverá ser prevista recorrendo a metodologias quantitativas de Regressão linear e não linear. Frequentemente, as abordagens qualitativas são consideradas como as mais adequadas nestes contextos de maior variabilidade e incerteza sobre a evolução dos dados. O autor refere que tem sido demonstrado que os modelos de séries temporais e as previsões qualitativas são ineficientes quando se pretende a previsão de observações atípicas:

Quantitative methods are very sucessful in producing baseline forecasts of time series; however these models fail to forecast neither the timming nor the impact of special events ... Forecasters prefer to use their own judgment for adjusting for forthcoming special events, but humans' efficiency in such tasks have been found to be deficient. ... As far the forthcoming irregular events are concerned, the only interest is on the potential impact of such event on the baseline forecast. ... The search for effective modeling procedures is therefore likely to be a fertile research field for many years to come and the synergy required from the information tecnology is rather crucial (via *Data Mining* and Information Managment techniques and algorithms. (p.1, p.2, p.16)

Este autor avalia a capacidade predictiva da Regressão linear Múltipla, de Redes Neurais e do KNN relativamente ao acréscimo que se verifica nas audiências totais com a transmissão de determinados eventos desportivos. A base de dados de audimetria grega integra uma amostra de 46 programas desportivos do período que decorre desde 1996 a 2000. As Redes Neurais proporcionaram os melhores resultados e devem, segundo o autor, ser adoptadas em contextos de maior complexidade e de não linearidade dos dados.

As margens de erro, quantificadas com o Erro Absoluto Médio (EAM), apresentam o valor de 9,9 para as Redes Neurais e de 12.68, 12.89 e de 14.54 respectivamente para um modelo de 3KNN, KNN e Regressão linear Múltipla. Salienta a dificuldade de realizar previsões com dados reais de audiências televisivas e a necessidade das metodologias serem mais eficientes do ponto de vista computacional para serem aplicáveis às bases de dados que surgem no meio industrial. Os programas foram caracterizados por um perito em análise de audiências em função de três variáveis: a sua importância, competitividade e segmento horário. As análises de segmentação que foram realizadas às séries temporais são relevantes neste contexto e o canal de transmissão também deve ser considerado, uma vez que difere o perfil as características do seu público.

A área do *Data Mining*, certamente, que apresenta o maior potencial de aplicação neste



contexto. Sugerimos nomeadamente o “Temporal Reasoning” que viabiliza a identificação de Regras Temporais (Valente & Chavarrias, 2000; Guimaraes, 2000; Morchen & Ultsch, 2004) em dados indexados pelo factor temporal, descrevendo relações entre sequências ou eventos instantâneos.

Fig. VI.5.1 – Interação entre o desempenho de diversos suportes horários.

Fonte: Elaborado pela autora.

Estas metodologias podem ser utilizadas quando estão presentes complexos padrões temporais, como, por exemplo, tendências crescentes ou decrescentes em segmentos de séries temporais simultâneos ou adjacentes (*Fig. VI.5.1*). Devemos diferenciar as observações em função da dimensão e evolução da sua repercussão nas audiências dos canais concorrentes e nos suportes adjacentes. Um evento desportivo poderá não incutir uma grande alteração na exposição global ao *media*, mas, sim, ao nível dos *shares* e de *ratings*. Por outro lado, em determinados contextos pode-se verificar o inverso e as alterações em *Tvr%* e *Shr%* podem-se complementar e anular ao ponto de não influenciar a evolução dos *ratings*. A sua repercussão nos programas adjacentes dos diversos canais também apresenta uma grande variabilidade.

Uma vez identificadas as interações existentes nas ST multivariadas, a previsão das audiências dos suportes poderá corresponder a uma função da subamostra anteriormente identificada e no caso mais simples apenas à sua média. Segundo *Nikolopoulos (2010, p.16)* os algoritmos KNN devem ser adoptados: “Some more effort should be given on the improvement of the performance of Nearest Neighbour approaches as they are quite intuitively appealing.” Contudo, salienta que as suas margens de erro absoluto médio aproximam-se dos 13%.

De acordo com alguma bibliografia de economia desportiva, as variáveis explicativas da evolução das audiências, frequentemente, não correspondem a variáveis de previsão (*Reeth, 2011*), e, logo, as análises multivariadas causais não são uma boa opção no contexto predictivo. Tem sido observado uma considerável evolução das audiências no decorrer de um evento desportivo e que é função da “qualidade” do desempenho dos jogadores e do grau de incerteza do seu resultado, entre outras variáveis dificilmente quantificáveis e/ou previsíveis.

Consequentemente, devemos essencialmente tentar conseguir alguma informação quantitativa sobre a mais provável evolução dos principais indicadores nos diversos segmentos horários e que permitirá aos anunciantes realizar uma melhor negociação com as estações televisivas, no caso de optarem por esta tipologia de programas.

“On the expectation of a future irregular event, an expert should adjust his statistical forecasts ... There is a need for automated support from FSS so as to enhance this adjustment process.” (*Nikolopoulos, 2010, p.5*). Frequentemente, os planificadores preferem investir em conteúdos com menores *ratings* para ser viável alcançar os seus objectivos de frequência em função do orçamento disponível, ao invés de tentar alcançar de forma pouco efectiva a totalidade do alvo.

If there's a choice to place one spot in a highly rated program or to place several spots in lower-rated programs with the same number of gross rating points, all *media* directors would prefer the latter even though placing several spots in lower-rated programs leads to lower reach. Despite Jones call for reach, the *media* directors universally believe in frequency. (*Makienko, 2012, p. 61*)

VI.6. OS MODELOS DE EXPOSIÇÃO QUANDO NOVOS CONTEÚDOS SÃO INTRODUZIDOS NUMA GRELHA DE PROGRAMAS

No *Gráfico IV.1.1* e no *Gráfico IV.1.2* constam algumas metodologias que viabilizam a previsão em contextos onde não se dispõe de dados suficientemente objectivos. Destacam-se as metodologias Delphi, Juízos de Valor, Análises Conjuntas, Simulações e Analogias Históricas, e por conseguinte incide maioritariamente nas análises qualitativas.

Podemos adicionar às metodologias enumeradas, a Teoria dos Jogos, as Intenções e Expectativas, os Sistemas Expertos, a Decomposição e a Previsão de Mercado. A previsão das audiências para novos programas será mais difícil de alcançar, mas, por outro lado, dependerá das tipologias e faixas horárias em análise, assim como da variabilidade dos padrões de consumo televisivo do segmento alvo e das variáveis que os condicionam.

Uma das possíveis perspectivas será recorrendo aos modelos de *ratings*, sendo esta a alternativa que minimiza os custos e o tempo a investir no seu desenvolvimento. A análise da evolução temporal do desempenho dos suportes horários em alguns contextos onde se observa uma maior estabilidade poderá ser suficiente.

Em alternativa, podemos considerar os valores históricos dos programas mais semelhantes ao programa em análise e transmitidos nas mesmas faixas horárias ou próximas, no mesmo canal e com um semelhante perfil do público-alvo.

Uma vez que a maior parte do painel não tem TV Cabo, os telespectadores encontram-se muito limitados ao nível das suas possibilidades de selecção, e, logo, esta perspectiva de análise poderá conduzir a resultados satisfatórios. Recorrendo a análises de segmentação das séries temporais podemos identificar os segmentos horários e as amostras a considerar nas análises. Quando se recorre aos cenários históricos pode na realidade dar-se uma grande variabilidade de valores.

Se um canal transmite um programa com um maior índice de exposição e introduz-se um segundo programa da mesma tipologia nesse período horário, a previsão do seu valor para um novo programa poderá estar baseada em valores médios superior ao real. Tudo dependerá do grau de repercussão das variáveis estruturais no segmento de telespectadores em análise. Os telespectadores devem considerados como activos em ambientes estruturais, ou seja, nem sempre os telespectadores apresentam uma total liberdade para que as selecções expressem as suas preferências. A frequente escolha da melhor alternativa ou de uma escolha estrutural poderá implicar uma menor variabilidade dos dados.

Na realidade, quanto mais os padrões de consumo televisivo dependerem da influência dos factores estruturais, menores podem ser as alterações decorrentes das substituições dos conteúdos, ou serem mais facilmente previsíveis e a abordagem anteriormente descrita de *Data Mining* Sequencial poderá ser considerada ou os modelos de *ratings* em função da evolução dos dados para os suportes horários em análise. Uma vez que grande parte do painel não tem TV Cabo e quando impera a experiência televisiva, as alterações de conteúdos podem não incutir uma forte variabilidade ao consumo televisivo. Também, verificámos que não existe uma correspondência directa entre a variabilidade dos fluxos de audiências e os valores de *ratings*.

Neste sector, quando impera a variabilidade dos dados, considera-se como necessário recorrer em maior grau a abordagens qualitativas, entre as quais, também, se destacam as sondagens de intenção de ver, enquanto não se dispõe de um conjunto de observações para esses novos suportes que permitam equacionar outras perspectivas de análise. Contudo, uma maior compreensão dos factores que determinam a exposição permite prever as tendências mais prováveis na evolução dos dados de forma mais objectiva.

A abordagem de *Data Mining* não proposicional proposta com a contínua actualização das sequências mais características deve ser considerada. Apenas com uma análise dos dados históricos podemos verificar se em determinado segmento imperam alguns dos comportamentos estruturais, o que condicionará as abordagens metodológicas mais adequadas e a variabilidade dos dados. Se existir uma maior actividade no processo de selecção e a ausência de regularidade do ponto de vista agregado, as abordagens qualitativas são necessárias e podem-se desenvolver projecções estocásticas.

A nível nacional, verificamos a existência de uma certa fidelização das audiências face a determinados suportes se consideramos determinadas faixas horárias e canais. Por conseguinte, a substituição de um programa por outro da mesma tipologia como se verifica no período da noite, poderá não levar a rápidas alterações nas selecções realizadas, especialmente se os lares não tiverem TV Cabo.

Nesses contextos, a previsão poderá se resumir à procura dos programas mais semelhantes e próximos nas faixas horárias. Como se verificou a estrutura horária é determinante para a previsão dos indicadores de audimetria e verifica-se uma forte influência dos factores de disponibilidade que determinam a intensidade dos comportamentos estruturais. Se existir estabilidade ao nível da evolução histórica dos *ratings* as análises univariadas e as análises multivariadas podem ser consideradas.

Quando a nova programação está associada a uma nova tipologia de programas, e, logo, falta um conjunto completo de observações históricas, a previsão das audiências desses novos programas terá que se basear principalmente nos índices de programas e

tipologias mais semelhantes e transmitidos inclusive nos restantes canais. Isto pode supor um erro se a audiência total dentro dessa parte do dia for diferente da histórica. Verificámos que a exposição global ao *media* apresenta uma evolução exponencial ou polinomial no decorrer da maior parte da emissão diária. A evolução dos *ratings* históricos dos suportes horários também condicionará a viabilidade de recorrer aos modelos de *ratings*. Outras abordagens de previsão qualitativa também correspondem a uma perspectiva de análise que *a priori* se poderia pensar como necessária nestes contextos, como, por exemplo, as sondagens ou recorrer a painéis de opinião.

As audiências dos jogos de futebol reflectem a existência de um grande segmento de telespectadores que são mais activos nas suas selecções e atribuem relevância aos conteúdos e que podem contribuir para uma maior alteração na evolução dos indicadores e no desempenho dos diversos canais. A principal fonte de variabilidade e incerteza será a sua capacidade de mobilização de telespectadores com uma exposição muito irregular e mais pontual.

Uma nova tipologia poderá apresentar potencial de mobilização de novos telespectadores, e, não, apenas incutir uma redistribuição ao nível dos segmentos que apresentam um maior consumo televisivo. Contudo, uma vez que a exposição pode não ser intencional, ou quando o é dificilmente será conhecida algumas semanas antes de ocorrer, as sondagens de intenção de ver podem produzir inevitavelmente resultados insatisfatórios e o analista deverá corrigir esses resultados em função do seu conhecimento do mercado. Nas sondagens o procedimento consiste em enviar por correio um questionário simples e breve, assim como uma carta para o seu reenvio, mas, por vezes, as respostas podem ser recolhidas telefonicamente. Estas decisões estão obviamente dependentes dos factores económicos das empresas que medem as audiências televisivas. No questionário, os telespectadores indicam quais são as faixas horárias que vão ver e os canais e ainda que inconscientemente acabam por expressar supostas preferências e não têm em atenção as estruturas que as restringem.

Será, eventualmente, mais aconselhável recorrer a painéis, porém, também, não proporcionam informação que é necessária à construção dos planos de *media*, sendo mais relevantes para a construção das grelhas de programas. A abordagem de *Data Mining* não proposicional também poderá ser equacionada nestes contextos. Qualquer uma das restantes abordagens também apresenta significativas limitações.

As Ciências Sociais não são uma Ciência Exacta, e, logo, as análises vão sempre apresentar margens de erro que podem atingir valores consideráveis. As previsões consideradas viáveis nas ciências não duras são apenas direccionais e inexactas. Apesar da dificuldade em conseguir alcançar resultados razoavelmente precisos, inevitavelmente em todos estes contextos novas abordagens metodológicas devem ser identificadas, sendo talvez a área do *Data Mining*, uma das mais relevantes no futuro. Actualmente, existem abordagens analíticas que permitem tentar antever a ocorrência de observações pouco frequentes que surgem em outros domínios industriais.

CAPÍTULO VII

CONCLUSÕES E CONSIDERAÇÕES FINAIS

VII.1. CONCLUSÕES FINAIS

O principal objectivo das análises realizadas traduziu-se na identificação das perspectivas de análise mais adequadas à construção dos modelos de exposição, que servem de suporte à actividade de planeamento de *media*, e que incidem na quantificação das audiências televisivas, numa perspectiva explicativa e de previsão da evolução dos dados. A questão e subquestões colocadas no *Capítulo III* foram resolvidas, como se demonstra nos *Capítulos V* e *VI*. Perante a escassa bibliografia académica nacional, ao nível dos diversos modelos de *media*, a exaustiva caracterização da evolução dos dados de audimetria permitiu a identificação das melhores perspectivas de análise na óptica dos modelos de *ratings* e a fundamentação de determinadas propostas metodológicas alternativas.

Os modelos de exposição devem permitir prever as audiências televisivas passíveis de ser quantificadas através de diversos indicadores, e, em alternativa, contribuir para a selecção dos suportes publicitários mais adequados ao público-alvo de uma campanha publicitária. Estes modelos viabilizam, assim, a optimização do investimento financeiro realizado pelos anunciantes e a possibilidade de gerar maiores receitas para as estações televisivas. Na bibliografia da especialidade está demonstrada a inviabilidade de construção dos modelos de exposição sem recurso aos dados de audimetria, que podem ser classificados em indicadores agregados e acumulados.

Os indicadores acumulados permitem a interpretação da evolução dos *ratings*, a avaliação da regularidade dos comportamentos de consumo televisivo e da viabilidade da sua previsão, assim como a posterior identificação de perspectivas de análise que possam constituir uma alternativa aos modelos baseados na análise das taxas médias de audiências. Se não existirem regularidades nos dados, a indução e a previsão não serão viáveis. Estes indicadores viabilizam, também, a análise dos comportamentos “estruturais” de consumo televisivo, nomeadamente se estão presentes no painel de audimetria nacional, respectivas interações e a sua repercussão na evolução dos *ratings*. Por outro lado, os indicadores agregados, também designados de “*gross measures*” de audimetria, condicionam o custo de um plano de *media*, uma vez que permitem o cálculo das audiências acumuladas, para além de serem determinantes para a aquisição de inserções no “*Spot Market*”.

Verificámos que a Análise Simbólica, e, em especial, diversas áreas de *Data Mining*, permitem que melhores resultados possam ser alcançados nas diversas etapas que integram a construção de uma campanha publicitária e que são função dos contextos, os quais foram previamente identificados na evolução dos dados. Estes resultados apresentam, nomeadamente, potencialidades ao nível da pré-selecção dos suportes publicitários e da segmentação dos dados, mas também ao nível da sua previsão e construção dos modelos de exposição, que antecedem a aplicação das análises de optimização.

Contudo, nas Ciências Sociais, as análises de previsão apresentam características que as diferenciam das análises que podem ser realizadas nas ciências exactas. Os sistemas humanos não devem ser considerados como sendo equivalentes aos sistemas físicos, onde está ausente a multidimensionalidade e a irregularidade que caracteriza os comportamentos sociais. Destacamos, em especial, o contributo do *Data Mining* de Séries Temporais, DM Sequencial e DM Multirrelacional, para além das metodologias proposicionais de Regressão não linear, entre as quais algumas Redes Neurais ou as Árvores de Regressão e os Instance Based Learners (IBL).

Os modelos de *ratings* apresentam uma significativa capacidade explicativa da evolução dos dados. No contexto programático, podem ser alcançados valores de percentagens de variância explicada no intervalo 80% - 90% e, para os *ratings* comerciais, no intervalo 75% - 90%, sendo inferior aos 50% para o indicador que quantifica o decréscimo relativo das audiências comerciais face aos correspondentes valores programáticos e representado pelo indicador *Spot_Ratio%*. Os melhores resultados derivam da aplicação das metodologias de Regressão não lineares e por recurso a metodologias de Regressão simbólicas.

O grau de precisão das análises de previsão é mais insatisfatório, quantificado com os valores dos coeficientes de determinação, erros quadráticos médios relativos e com as medidas de diversidade das Árvores. Os modelos de *ratings* podem ser construídos com recurso a análises univariadas ou multivariadas, mas considerando apenas as variáveis conhecidas ou que podem ser calculadas no momento de realizar uma previsão. Apesar das inúmeras variáveis que foram calculadas, uma vez que a informação que consta das bases de audimetria é insuficiente, não foi viável identificar um bom conjunto de variáveis de previsão no contexto programático.

Com base nas análises univariadas e multivariadas realizadas, verificámos que os modelos de *ratings* programáticos podem constituir uma opção do ponto de vista da previsão apenas para determinadas faixas horárias e segmentos sociodemográficos ou tipologias de programas que apresentam uma maior estabilidade, considerando, por exemplo, a tarde e a manhã ou dados do fim-de-semana, desde que não sejam transmitidos determinados programas desportivos. As análises programáticas devem ser realizadas para os suportes temporais e, posteriormente, os seus valores extrapolados para suportes programáticos. A segmentação e o *clustering* das séries temporais de *ratings* poderá ser considerado, assim como os modelos de Regressão linear simples ou de Regressão não linear, considerando a sua relação com a exposição global ao *media*. Os valores de *shares* não constituem uma variável com capacidade predictiva devido à sua maior irregularidade.

A previsão univariada das taxas médias de audiências para o total da TV, recorrendo a algoritmos de IBL ou de Regressão local, não apresenta, também, resultados muito satisfatórios, ficando aquém dos 80%. Contudo, estes valores são superiores aos valores que é possível encontrar na bibliografia internacional da especialidade. Por outro lado, os resultados na área comercial podem ser mais precisos. Constatámos a existência de uma forte associação entre os valores de *ratings* programáticos e os valores de *ratings* comerciais, sendo os factores de planeamento determinantes para o desempenho das diversas inserções transmitidas no decorrer de um intervalo publicitário.

Por vezes, os dados de *ratings* não viabilizam a construção dos modelos de exposição com um grau de precisão que possa ser considerado satisfatório e, logo, devem ser identificadas alternativas. Nas análises realizadas, privilegiou-se uma maior compreensão da influência estrutural nos fluxos de audiências e na evolução dos *ratings*, uma vez que podem apresentar uma maior capacidade predictiva e utilidade na explicação do comportamento agregado. Verificámos que, a nível nacional, deverá existir uma maior influência estrutural no consumo televisivo, mas, no geral, notámos uma reduzida regularidade dos comportamentos, constatada através dos resultados das análises de duplicação das audiências.

No modelo integrado e explicativo da exposição televisiva, apresentado na sequência das conclusões retiradas sobre a análise dos comportamentos “estruturais”, consta a identificação de diversos cenários de exposição intencional e não intencional. Podemos identificar contextos de exposição intencional de lealdade programática ou de exposição não intencional neste mesmo contexto, imperando a experiência televisiva. A exposição poderá, também, ser não intencional por efeitos de herança ou por lealdade ao canal, ou intencional por lealdade tipológica. Por outro lado, os conteúdos podem, também, ser evitados por decisões conscientes dos telespectadores ou como consequência de variáveis externas que os inviabilizam. A duplicação de audiências entre programas próximos ou adjacentes não resulta apenas da inércia do público, mas, sim, também da interacção existente entre os comportamentos de lealdade programática, tipológica e ao canal.

Assim, ficam demonstradas as limitações de determinadas metodologias de previsão das audiências adoptadas pelo sector para a análise dos dados que apresentam uma maior instabilidade dos valores de *ratings*, inclusive após o desenvolvimento da audimetria. As sondagens de intenção de ver, entre outras abordagens que visam a quantificação das audiências privilegiando a análise das preferências e interesses do painel, produzem resultados muito insatisfatórios do ponto de vista da previsão, mas também explicativo. As restantes metodologias qualitativas não permitem, também, ao sector dispor de informação que é essencial para o planeamento de *media*.

Devem ser identificados modelos de previsão quantitativos realmente adequados aos contextos de maior instabilidade dos dados, e tendo em atenção o carácter imprevisível e indeterminado dos comportamentos individuais. Será necessário considerar o colectivo e devemos identificar, nos dados, a estrutura dos seus comportamentos, a qual poderá permanecer mais estável no decorrer do tempo. Atendendo às múltiplas variáveis não quantificáveis e previsíveis que influenciam os comportamentos, a construção de modelos causais multivariados não constitui uma boa opção.

Para os segmentos que apresentam uma menor regularidade do consumo televisivo, as metodologias factoriais simbólicas podem contribuir para que lhes possam ser associados determinados conteúdos. Deste modo, atendendo à sua elevada dimensionalidade, conseguimos restringir o número de suportes a considerar nas etapas posteriores. Após uma eventual segmentação da base de dados através dos algoritmos de *clustering* de sequências e promovendo, paralelamente, a redução da dimensionalidade dos suportes, a construção de Regras sequenciais de *data mining* ou de associação permite a identificação dos suportes que devem ser incluídos nos planos de *media*, desde que visando os segmentos onde se torne mais evidente o desenvolvimento de hábitos de consumo televisivo. Uma vez que a grelha e os comportamentos evoluem no decorrer do tempo, devem ser consideradas as Regras para ambientes dinâmicos, para que se verifique uma contínua actualização dos seus resultados.

As análises devem ser aprofundadas para os segmentos mais instáveis através da construção de Regras de classificação ou Regressão multirrelacionais e por utilização de Redes Bayesianas que tenham em atenção as perspectivas explicativas da exposição televisiva que caracterizem, em maior grau, o painel nacional. Nos segmentos de telespectadores que apresentam uma exposição mais regular, podemos, também, considerar as sequências mais características que permitem identificar os suportes a contratar, assim como recorrer a modelos Logit ou a Regras de classificação de Bayes, entre outras abordagens.

A previsão da repercussão da transmissão de eventos desportivos corresponde à tipologia de programas que poderá inculir uma maior variabilidade aos dados. Neste contexto, sugere-se o recurso a Regras temporais que tenham em atenção as eventuais interacções observadas em subsequências das séries temporais multivariadas de valores de *ratings*, com eventual posterior aplicação do MBL. Devemos dispôr de informação qualitativa a respeito da futura evolução mais provável dos dados que permita aos anunciantes negociar com as estações televisivas.

Quando se verificam consideráveis alterações nas grelhas de programas, as análises não devem, também, ser restringidas apenas à previsão qualitativa. Verificámos que prevalece a influência estrutural e, logo, as alterações nem sempre se repercutem de forma muito significativa na evolução dos dados. Os padrões de consumo televisivo podem apresentar uma evolução mais gradual e a exposição estará em maior grau enquadrada nas sequências mais características de determinado segmento ou telespectador. Em especial para os segmentos onde seja superior a influência das variáveis individuais descritas na perspectiva de “gratificações e usos do *media*”, será necessário avaliar novas Regras de classificação multirelacionais. Inclusive, na perspectiva dos modelos de *ratings*, verificámos que alguns canais, nomeadamente os estatais generalistas, apresentam uma maior estabilidade para este indicador, apesar de serem frequentes as alterações na estruturação das suas grelhas. Portanto, os modelos de *ratings* podem, também, ser considerados, inclusive, nas faixas horárias que apresentam uma maior exposição ao painel.

Cada contexto em análise requer que sejam adoptadas diferentes perspectivas e as análises de previsão nas Ciências Sociais não podem apresentar o grau de precisão que, por vezes, é viável nas ciências exactas, onde estão em análise sistemas físicos que não apresentam a multidimensionalidade que caracteriza os sistemas humanos. Por conseguinte, devem ser desenvolvidas perspectivas de análise realmente adequadas às características e evolução dos dados de audimetria. O desenvolvimento de determinadas áreas de *Data Mining* permitirá que melhores resultados possam ser alcançados na área do planeamento de *media*, assim como a sua integração com a investigação qualitativa e com fontes de dados alternativas que devem complementar a informação obtida através da tecnologia de audimetria ou qualquer outra tecnologia que lhe seja equivalente.

VII.2. SUMÁRIO DE CONTRIBUIÇÕES

Os modelos de exposição integram uma das principais etapas necessárias à construção dos modelos matemáticos que viabilizam a aquisição de publicidade. Num domínio, frequentemente, classificado como acausal, imprevisível e indeterminado, as análises de previsão e as análises explicativas dos comportamentos de consumo televisivo e dos indicadores que deles resultam apresentam grandes interdependências, sendo essenciais para a actividade de planeamento das campanhas publicitárias.

Verificámos que a bibliografia académica internacional existente nesta área é reduzida, sendo-o ainda mais a nível nacional, existindo inúmeras limitações a diversos níveis. A dificuldade de acesso aos dados e a inadequabilidade ou restrições que as metodologias existentes colocam quando se pretende a análise das bases de audimetria são apenas alguns dos factores que contribuem para esta realidade. Por exemplo, ao nível dos modelos de exposição, os trabalhos incidem em dados anglo-saxónicos de *ratings* anteriores ao ano de 2000. Prevalece a análise dos dados das faixas horárias da noite e em maior grau apenas para o universo televisivo. Do ponto de vista metodológico, são muito pontuais os contributos ao nível das análises não lineares de *Data Mining* e não foi possível identificar quaisquer análises simbólicas de análise de *ratings*.

Atendendo à escassa informação de caracterização dos suportes, que consta das bases de audimetria, o número de variáveis considerado nas análises explicativas multivariadas que é possível encontrar na bibliografia da especialidade é insuficiente para conseguir compreender a evolução dos dados e para viabilizar a construção de modelos que possam apresentar alguma capacidade predictiva. É, também, muito pontual a análise dos ficheiros base de audimetria e as propostas teóricas ou validadas empiricamente ao nível das abordagens alternativas aos modelos de *ratings*.

Verificámos que os dados em que incidem as análises de duplicação das audiências nem sempre são os mais adequados, pela dificuldade em aceder e construir as bases de dados que viabilizam uma análise mais objectiva destes comportamentos. As conclusões alcançadas em determinadas áreas são influenciadas pela inadequabilidade das variáveis e metodologias utilizadas, como, por exemplo, por se verificar uma excessiva ênfase na utilização das abordagens de Regressão linear que apenas permitem avaliar a linearidade das relações entre as variáveis. Frequentemente, conclui-se que existe ausência de relação entre as variáveis quando ela poderá existir, mas não ser do tipo linear. Ao nível da previsão, os modelos nem sempre são validados correctamente.

O grau de maturidade que caracteriza determinada área apresenta os fundamentos da sua previsão e, nesta área, observamos que existem inúmeras lacunas na investigação. Quando é necessário tomar determinadas decisões, que podem estar motivadas, ou não, pela necessidade de prever ou de controlar um sistema que se encontra constituído por complexos componentes correlacionados, existe a necessidade de, previamente, analisar esse mesmo sistema. Quanto melhor se entender essa complexidade, melhor será a decisão, e, neste contexto, a selecção das perspectivas de previsão das audiências televisivas para o curto prazo.

As melhores abordagens dependem dos contextos e são função da regularidade e acessibilidade dos dados, assim como dos custos, do investimento temporal e da precisão dos resultados alcançados.

Consequentemente, a perspectiva a privilegiar deverá ser o desenvolvimento dos modelos de *ratings* e, em função dos resultados, devem ser avaliadas alternativas que apresentem uma crescente complexidade ao nível dos diversos recursos envolvidos, e que foram descritas no capítulo anterior. Os resultados das análises explicativas realizadas aos indicadores agregados e acumulados de audimetria permitiram avaliar a regularidade dos indicadores e dos comportamentos que os originam. A estabilidade dos dados condiciona a viabilidade das análises de previsão e a componente analítica a adoptar, sendo esta função dos contextos que foram previamente identificados na evolução dos dados.

Nas análises realizadas, privilegiou-se uma maior compreensão da influência estrutural nos fluxos de audiências e na evolução dos *ratings*, uma vez que os comportamentos estruturais podem apresentar uma maior capacidade predictiva e utilidade na explicação da evolução dos fluxos de audiências. Verificámos que prevalece a irregularidade da exposição ao *media*, mas existe uma maior estabilidade ao nível das selecções. Os modelos explicativos estruturais e integrados da exposição televisiva descritos pretendem demonstrar os múltiplos contextos de exposição não intencional ou da sua ausência, e que fundamentam alguns dos modelos de *Data Mining* apresentados, constituindo uma alternativa à análise dos *ratings*.

A componente metodológica determina a informação que poderá ser extraída das bases de dados de audimetria, sendo esta essencial para a construção dos modelos quantitativos de *media*, uma vez que viabilizam uma análise diacrónica da evolução dos dados em contextos onde não impera a regularidade dos comportamentos, mas, sim, a sua multidimensionalidade e acausalidade. Inúmeros autores referem que os instrumentos disponíveis previamente ao desenvolvimento da audimetria não viabilizam a construção dos modelos que são necessários à aquisição de espaço comercial, uma vez que a interpretação frequentista da probabilidade não se ajusta aos comportamentos sociais e não prevalece a regularidade dos dados.

As perspectivas qualitativas de previsão de dados são consideradas como insuficientes para o planeamento de *media*, sendo utilizadas, essencialmente, por parte das estações televisivas quando se pretende a construção das grelhas de programas incluindo novos programas e tipologias. Consequentemente, a fonte de dados privilegiada teria que ser a audimetria.

Os modelos explicativos da exposição televisiva podem ser agregados em duas grandes áreas que adoptam diferentes fontes de dados e de metodologias, e que perseguem objectivos contraditórios, nomeadamente ao nível da concepção do telespectador, enquanto agente activo ou passivo nas suas decisões de acesso ao *media* e de selecção de conteúdos.

Uma vez que a reduzida investigação internacional ao nível dos modelos explicativos que integram ambas as perspectivas demonstra que prevalece a influência estrutural, procurou-se aprofundar a análise dos comportamentos estruturais, avaliando em que medida estes podem ser identificados no painel nacional, qual a sua interacção e quais os factores que os condicionam, assim como a sua influência na evolução dos principais indicadores de audimetria que condicionam a venda do espaço comercial. A perspectiva estrutural corresponde à que, internacionalmente, recorre aos sistemas de informação para caracterizar a exposição televisiva. Os comportamentos estruturais resultam da estruturação das grelhas de programas.

Verificámos que, a nível nacional, existe uma maior influência estrutural, mas uma reduzida regularidade dos comportamentos, o que torna imprescindível o desenvolvimento de modelos de *ratings* ou a identificação de perspectivas alternativas de análise adequadas à evolução dos fluxos de audiências entre os suportes que são seleccionados em função de determinados critérios quantitativos e qualitativos. No modelo integrado e explicativo da exposição televisiva apresentado constam diversos cenários de exposição intencional ou não intencional, que resultam das influências estruturais e de variáveis individuais.

Assim, podemos identificar contextos de exposição intencional de lealdade programática ou não intencional, imperando a experiência televisiva, a exposição não intencional por efeitos de herança e de lealdade ao canal e a selecção intencional de lealdade tipológica. Os conteúdos podem, também, ser evitados de forma consciente ou a ausência de exposição poderá ser uma consequência de variáveis externas que a inviabilizam.

A duplicação de audiências entre programas próximos ou adjacentes não resulta, igualmente, apenas da inércia do público, mas, também, da interacção entre a lealdade programática, tipológica e ao canal. Consequentemente, as sondagens de intenção de ver, entre outras abordagens que visam a quantificação das audiências através de uma mera análise das preferências e dos interesses do painel, produz resultados muito insatisfatórios do ponto de vista da previsão, mas também explicativo.

Em termos globais, a constatada ausência de regularidade dos fluxos de audiências entre os suportes levanta a necessidade de avaliar quais são as potencialidades de áreas mais recentes e adequadas, e, em especial, do *Data Mining* e da Análise Simbólica. Verificámos que, após uma eventual prévia segmentação simbólica da emissão diária em função da variabilidade dos dados, é possível identificar os suportes a incluir em planos de *media*, recorrendo ao *Data Mining* Temporal, nomeadamente a Regras de Associação ou Sequenciais, devendo privilegiar-se a exposição global ao *media*.

Após a segmentação da base de dados através de algoritmos de *clustering* de sequências e promovendo a redução da dimensionalidade dos suportes, a construção de Regras temporais permite a identificação dos suportes que devem ser incluídos em planos de *media*, desde que visando segmentos onde se torne mais evidente o desenvolvimento de hábitos de consumo televisivo. O conceito de hábito nem sempre implica uma forte regularidade, discutindo-se este conceito que tem sido associado à perspectiva estrutural explicativa da exposição televisiva. Uma vez que a grelha de programas e o público evoluem no decorrer do tempo, as Regras para ambientes dinâmicos devem ser consideradas, de forma a verificar-se uma contínua actualização dos resultados.

A redução da dimensionalidade dos suportes a considerar nas bases de dados de exposição pode ser realizada recorrendo a metodologias simbólicas MDS quando estamos perante suportes que são transmitidos em séries ou a análises de Correspondência, MDS e à RL como alternativa à análise de *Clusters* e Discriminante linear nos contextos de uma única transmissão. Contudo, verificámos que existe uma considerável dificuldade de associação de determinados conteúdos aos segmentos sociodemográficos, como consequência da irregularidade da exposição e das influências estruturais. Quando os segmentos em análise apresentam uma menor regularidade, as análises devem ser complementadas com a aplicação de algoritmos Multirrelacionais e das Redes Bayesianas que permitam avaliar e identificar os factores com maior repercussão na exposição televisiva e com capacidade de previsão dos comportamentos considerando o curto prazo. As abordagens integradas devem ser privilegiadas, isto é, as que consideram variáveis enquadradas em diversas perspectivas explicativas da exposição.

O plano final será, também, o resultado das restrições orçamentais dos anunciantes, entre outras, que podem ser consideradas nos algoritmos de optimização e onde se pretende maximizar a cobertura do mercado-alvo e com valores de frequências que são função das características da campanha publicitária, mas que devem ter em atenção a complementaridade dos suportes e os padrões de consumo mais característicos dos segmentos em análise. Deve ser privilegiada a frequência estrutural de determinado segmento, uma vez que, do ponto de vista explicativo, não está demonstrada a existência de um valor que determine o sucesso ou insucesso de uma campanha, mas, sim, de múltiplos valores que dependem das características da campanha e que devem ser definidos pelos planificadores, apresentando subjectividade.

Atendendo às características que os sistemas de previsão devem apresentar, nomeadamente ao nível do custo, rapidez e acessibilidade dos dados, a construção dos modelos de exposição pode, também, ser viabilizada através da construção dos modelos de *ratings* que foram, igualmente, considerados e em maior grau a nível explicativo. Os comportamentos estruturais foram relacionados com a evolução dos *ratings*.

Considerou-se, como suportes, as faixas horárias, os programas, os espaços comerciais e as inserções comerciais e, como variáveis dependentes, os valores de *ratings*, uma vez que correspondem ao indicador que condiciona a venda do espaço comercial, assim como as taxas médias de audiências para o total da TV, os *shares* de audiências e o índice *Spot_Ratio%*, que quantifica a inter-relação observada entre as audiências programáticas e os valores comerciais. Recorrendo a determinadas abordagens lineares e não lineares enquadradas na Análise Simbólica, no *Data Mining* e na Estatística Multivariada, foi possível avaliar os contextos de maior e menor regularidade dos dados, bem como alguns dos factores responsáveis pela sua evolução. Destacamos, como principal fonte de variabilidade dos dados, algumas tipologias de programas em contra-programação ou na programação adjacente, o canal, as alterações horárias e, ao nível da duração dos suportes, as variáveis temporais que condicionam o perfil do público disponível e com predisposição para ver TV, assim como outras variáveis relacionadas com a estruturação da grelha. As variáveis de planeamento apresentam uma maior capacidade explicativa e predictiva face às que caracterizam os conteúdos.

Os modelos de *ratings* apresentam uma capacidade explicativa com valores no intervalo 80% - 90% para o contexto programático e no intervalo 75% - 90% para os *ratings* comerciais. Os *ratings* dos suportes horários são mais estáveis face aos valores programáticos, assim como a exposição global ao *media*, a qual apresenta uma considerável correlação e associação não linear com os valores dos *ratings*. A variável *share* corresponde a um dos indicadores que apresentam uma maior variabilidade, não constituindo uma variável de previsão para o período em análise. Os valores dos *ratings* comerciais são mais estáveis face aos valores programáticos que lhes estão associados, sendo elevada a associação entre ambos os indicadores e estando a sua evolução condicionada por parte dos factores de planeamento. Por exemplo, considerando os *ratings* programáticos, e por aplicação do algoritmo CHAID, alcançamos 87% de variância explicada, sendo esta de 82% para o algoritmo CART, incidindo as variáveis, essencialmente, em factores temporais, tipologias de programas, o canal e a fidelidade programática, entre outros factores de planeamento.

A capacidade explicativa dos diversos modelos multivariados que foram avaliados pode ser considerada satisfatória, mas a sua capacidade predictiva fica consideravelmente aquém da explicativa, uma vez que as variáveis identificadas nem sempre são conhecidas ou previsíveis quando se pretende realizar análises de previsão. As análises de previsão realizadas com Redes Neurais MLP e RBF permitiram alcançar 82% e 80%, respectivamente, de observações correctamente classificadas quando discretizadas em intervalos de 4% de amplitude. Foi necessário avaliar a associação dos *ratings* com os restantes indicadores que apresentam uma maior estabilidade e que, eventualmente, possam ser mais facilmente previsíveis.

Numa perspectiva explicativa da evolução dos dados, verificámos que as audiências comerciais dependem fortemente dos valores das audiências programáticas, se considerarmos os valores de *ratings*, mas, considerando os valores de *Spot_Ratio%*, os resultados são insatisfatórios. O comportamento da audiência depende de um grande e diversificado conjunto de variáveis e, logo, será difícil explicar de forma directa a evolução dos dados, ainda que recorrendo a um extenso número de variáveis e determinando, quer a sua influência individual, quer a conjunta.

No que se refere aos valores de *ratings*, verificámos que a sua variabilidade é substancialmente inferior à dos valores programáticos que lhes estão associados e que os modelos lineares e não lineares podem ser considerados em análises de previsão do indicador comercial, desde que quantificadas as audiências programáticas que apresentam uma maior variabilidade face às análises que incidem em suportes horários. As análises multivariadas devem ser consideradas quando os factores a avaliar são conhecidos ou são mais fáceis de prever do que a variável dependente.

As análises univariadas devem, também, ser avaliadas para a análise dos principais indicadores de audimetria e, em especial, para os *ratings*.

As análises univariadas realizadas com metodologias mais tradicionais, sendo estritamente extrapolativas, limitam a sua utilização directa a ambientes mais estáveis, mas podem ser utilizadas para a previsão de factores a utilizar em modelos causais e para o estabelecimento de tendências. Consequentemente, atendendo à associação observada entre os valores de *ratings* e os valores das taxas médias de audiências para o total da TV, este último indicador foi alvo de análises de previsão univariadas, privilegiando-se algumas abordagens alternativas aos métodos de decomposição ou *Box Jenkins*.

As análises univariadas de previsão para suportes programáticos e/ou horários apresentam margens de erro que são, na generalidade, consideráveis e, em especial, se considerarmos os dados de fim-de-semana e por aplicação de modelos locais de Regressão, como, por exemplo, os IBL, a Regressão *kernel*, a Regressão local linear e a Regressão local polinomial. A nível semanal, conseguiu-se, no máximo, uma redução do erro quadrático médio em 75,7% e de 58,7% ao fim-de-semana. A utilização das Médias Móveis e das Redes Neurais MLP não acrescentou mais-valias. Nas análises de segmentação simbólicas, as percentagens de variância explicada podem alcançar os 90%.

Numa perspectiva explicativa da evolução do índice *Tvr%*, verificamos que poderá apresentar alguma independência face aos conteúdos transmitidos. Apesar de os seus valores médios serem semelhantes, de ano para ano, esta variável poderá alcançar uma considerável variabilidade, ultrapassando os 24% se considerarmos os dados no decorrer de um ano. Nas análises realizadas, foram identificados os contextos em que a exposição global ao *media* é mais irregular.

Também numa perspectiva de análise univariada dos dados, a análise dos valores de *ratings* recorrendo a algoritmos de segmentação e *clustering* das séries temporais viabiliza resultados mais satisfatórios, em determinados contextos que apresentam uma menor variabilidade dos dados, especialmente para os períodos da manhã e tarde.

As previsões qualitativas, numa óptica de planeamento de *media*, apresentam uma maior relevância quando são transmitidos programas que podem gerar observações atípicas, sendo necessário, essencialmente, prever a evolução mais provável dos indicadores para se poder realizar uma boa negociação com a estação televisiva. Devemos dispor de alguma informação que permita justificar um menor investimento nesses suportes. Atendendo às análises realizadas, a previsão da repercussão da transmissão de eventos desportivos corresponde à tipologia de programas que poderá incutir uma maior variabilidade aos dados. Nestes contextos, sugere-se recorrer a Regras temporais que tenham em atenção as fortes interacções observadas nas séries temporais multivariadas, com posterior aplicação de MBL.

Verificámos que alguma bibliografia de economia desportiva, mais recente, apresenta variáveis explicativas da exposição que não são, *a priori*, quantificáveis e previsíveis, logo, as análises multivariadas dificilmente podem ser equacionadas nestes contextos. Uma vez que as análises quantitativas são essenciais, ainda que, posteriormente, complementadas com as qualitativas, será necessário, em primeiro lugar, conseguir prever a repercussão desta tipologia de dados nos indicadores que avaliam o desempenho dos canais concorrentes, inclusive para horizontes temporais subsequentes e antecedentes. A evolução mais provável dos dados poderá ser identificada através do *Data Mining* Temporal e a previsão poderá, eventualmente, ser realizada recorrendo a metodologias IBL que se baseiam no conceito da analogia histórica.

Quando se verificam consideráveis alterações nas grelhas, as análises não devem ser restringidas à previsão qualitativa. Nos modelos de análise dos fluxos de audiências, devem ser consideradas as Regras de associação para ambientes que não são estáticos. De acordo com as análises realizadas aos indicadores agregados e acumulados de audimetria, verificámos que prevalece a influência estrutural.

Logo, as alterações na grelha nem sempre se repercutem de forma muito significativa na evolução dos dados considerando o curto prazo, assim como, por vezes, também não ao nível da evolução dos *ratings*. Os padrões de consumo televisivo podem apresentar uma evolução mais gradual e a exposição poderá estar, em maior grau, enquadrada nas sequências mais características de um segmento, apesar de os telespectadores poderem variar as selecções de semana para semana. Em especial para os segmentos onde for superior a influência das variáveis motivacionais, entre outras variáveis individuais descritas na perspectiva de gratificações e usos do *media*, será necessário avaliar novas Regras de classificação multirrelacionais. Na perspectiva dos modelos de *ratings*, verificámos que os dados do canal RTP1 apresentam uma maior estabilidade de valores de *ratings*, apesar de serem mais frequentes as alterações na estruturação da sua grelha. Por outro lado, os canais privados generalistas apresentam uma evolução que caracteriza o período semanal *versus* os dias de fim-de-semana.

Em suma, são inúmeras as perspectivas de análise apresentadas que podem viabilizar a construção dos modelos de *media* que são essenciais para a aquisição de espaço comercial e que são função da regularidade dos indicadores e dos comportamentos. A sua selecção não dependerá, apenas, da precisão dos resultados, mas, também, da rapidez, custo e disponibilidade de dados. Cada utilizador deverá seleccionar a abordagem que considere apresentar uma melhor relação custo-benefício.

VII.3. LIMITAÇÕES DAS ANÁLISES REALIZADAS E NOVOS PROBLEMAS DE INVESTIGAÇÃO NA ÁREA DO PLANEAMENTO DE *MEDIA* E DOS MODELOS DE EXPOSIÇÃO

A evolução observada ao nível dos instrumentos metodológicos não tem acompanhado a complexidade do consumo televisivo, e apenas o desenvolvimento de novos métodos quantitativos permitirá a sua análise no futuro (Weber, 2002).

As análises realizadas e propostas tiveram como objectivo avaliar se a Análise Simbólica e determinadas metodologias de *Data Mining* permitem contribuir para que melhores resultados sejam alcançados numa área que é essencial para o sector televisivo e que tem colocado inúmeras dificuldades a quem pretende investir em publicidade televisiva. No *Capítulo VI* desta tese consta a discussão de eventuais contributos de determinadas áreas metodológicas para o planeamento de *media*.

A Análise Simbólica viabiliza uma melhor representação da informação existente nas bases de audimetria e o *Data Mining* permite a identificação de padrões, tendências e interacções existentes em bases de dados que apresentam uma elevada dimensionalidade e onde as variáveis se relacionam de forma não linear, como se verifica no contexto da audimetria. Não foi possível identificar, na bibliografia da especialidade, trabalhos de estatística aplicada ao nível da Análise Simbólica e apenas um número muito restrito de trabalhos de *Data Mining* e que não incidem nos contextos de análise que são mais relevantes para o planeamento de *media*.

Consequentemente, os novos problemas de investigação e as sugestões para futuras análises descritas no *Capítulo VI* incidem, essencialmente, em questões metodológicas. Tentou-se identificar e sugerir o maior número possível de perspectivas de análise que possam contribuir para a construção dos modelos de exposição, o que apenas foi viável por terem sido realizadas inúmeras análises descritivas e explicativas da evolução dos indicadores de audimetria que permitiram contribuir para uma maior compreensão da sua evolução. Não foi viável a implementação prática da totalidade das propostas que pretendem contribuir para a construção dos modelos de exposição, por inexistência de recursos e de bases de dados realmente adequadas, e atendendo às óbvias limitações temporais.

A reduzida bibliografia internacional e nacional existente nesta área, segundo inúmeros autores (*Danaher et al., 2011; Weber 2002*), resulta das dificuldades inerentes ao acesso aos dados de audimetria, detidos, em geral, por multinacionais, uma vez que estes apresentam um elevado valor comercial. A investigação é realizada, maioritariamente, no meio empresarial e os resultados alcançados não são divulgados. A ausência de bibliografia, o elevado investimento temporal, os recursos informáticos, a inadequabilidade das metodologias mais tradicionais de análise estatística de dados e o abrangente *know-how* requerido para a sua análise são, também, alguns factores que têm contribuído para esta tendência. A estrutura das bases de dados de audimetria a que as agências de meios têm acesso não facilita, certamente, a sua análise. A informação encontra-se demasiado dispersa e dificilmente se consegue inter-relacionar, para além da necessidade de corrigir e calcular um grande número de variáveis.

Em futuras análises a realizar, nomeadamente com os dados dos ficheiros base de audimetria, avaliar-se-á empiricamente algumas metodologias propostas na área do *Data Mining* não proposicional. Porém, estas encontram-se fundamentadas do ponto de vista teórico e resultam das conclusões alcançadas com as múltiplas análises que foram realizadas aos indicadores agregados e acumulados de audimetria que permitiram avaliar a regularidade dos indicadores e dos comportamentos que estão na sua base e a identificação de variáveis explicativas da sua evolução.

Apenas com o futuro desenvolvimento de determinadas áreas de DM poderão ser alcançados melhores resultados na área do planeamento de *media*, nomeadamente na área multirrelacional e temporal. Algumas das metodologias existentes apresentam consideráveis limitações quando se pretende a análise de bases de dados que apresentam uma elevada dimensionalidade, bem como lacunas e uma forte dispersão da informação ou quando prevalecem as variáveis numéricas e se pretende realizar análises de previsão. Os algoritmos de previsão e de segmentação devem ser computacionalmente mais eficientes para serem aplicáveis a bases de dados reais (*Nikolopoulos, 2010*).

Nomeadamente, certas metodologias de *Data Mining* não proposicional, como, por exemplo, na área do *Multirelational Data Mining*, apesar dos esforços desenvolvidos, adequam-se em maior grau à classificação de dados e ao tratamento de variáveis nominais. Os algoritmos proposicionais, apesar da sua simplicidade e eficiência, apresentam, como limitações, a sua restrita capacidade para representar conhecimento subjacente e relações complexas entre os dados.

A aprendizagem multirrelacional pretende ultrapassar estas condicionantes, mas tem sido considerada como computacionalmente ineficiente, apresentando consideráveis limitações quando é necessário considerar variáveis numéricas, sugerindo-se, nestes contextos, a proposionalização dos dados e a sua transformação em bases com um formato tabular (*Knobbe, Siebe & Van der Wallen, 1999*). Consideramos que a aplicação de metodologias simbólicas será mais adequada, em detrimento de uma agregação dos dados, isto é, uma síntese dos dados recorrendo a variáveis que representem a média e variância observada. A agregação dos dados provenientes de múltiplas relações pode resultar em perda de informação e conduzir a resultados menos claros.

As bases de audimetria apresentam dados demasiado dispersos e não inter-relacionados, e a construção de uma base de dados relacional, a partir dos dados disponíveis através do *software* Telereport e Publitex, exigiria um elevado investimento temporal e não acrescentaria, segundo as análises realizadas, grandes mais-valias ao nível dos modelos de *ratings*. As metodologias multirrelacionais devem ser mais relevantes para a construção de modelos que permitam a interpretação da estrutura dos comportamentos de um segmento de telespectadores.

O *Data Mining* Sequencial permite-nos a identificação dos suportes a incluir em planos de *media* e sobre os quais incidir as análises de previsão em etapas posteriores, para além de uma melhor caracterização da exposição televisiva.

As formas mais simples de Regras são as booleanas, unidimensionais e de um só nível, mas devem ser avaliadas as potencialidades de algoritmos que permitam implementar Regras quantitativas multidimensionais e multinível e que considerem intervalos de valores. O desenvolvimento das Regras sequenciais permitirá um melhor relacionamento da exposição verificada em diferentes sessões televisivas com diversos graus de desfasamento temporal. Deverá, também, ser aprofundada a integração da informação que resulta dos modelos de exposição e a construção do plano final recorrendo a algoritmos de optimização. Em cada campanha, terá que ser empreendida uma avaliação de quais os valores mais adequados de cobertura e de duplicação das audiências.

Para além das limitações que algumas áreas metodológicas e correspondentes algoritmos que são relevantes para o planeamento de *media* apresentam, encontramos, também, lacunas ao nível dos modelos explicativos do consumo televisivo. Estes modelos contribuem para o planeamento de *media*, numa perspectiva alternativa aos modelos de *ratings*, e devem ser aprofundados, promovendo a avaliação das interdependências observadas entre as variáveis associadas a diversas perspectivas teóricas explicativas do comportamento dos telespectadores.

Na presente tese, foi apresentado um modelo estrutural da exposição televisiva que permitiu avaliar as interacções observadas entre os comportamentos estruturais, assim como a sua repercussão na evolução dos *ratings*. Os inúmeros contextos de exposição não intencional ou de ausência de exposição influenciada por variáveis estruturais inviabilizam que possam ser alcançados bons resultados na área da previsão quando apenas se consideram as motivações e as preferências do painel. Por exemplo, em Wyndham (1994) as margens de erro podem alcançar os 30%. De acordo com a bibliografia internacional mais recente, e que incide em mercados que apresentam uma maior oferta televisiva face ao que se verifica actualmente a nível nacional, prevalece a influência estrutural e são reduzidas as interacções observadas com as variáveis associadas à perspectiva das gratificações e usos do *media*.

Por conseguinte, persistem inúmeras incógnitas nesta área e devem ser construídos modelos mais precisos a respeito dos factores que determinam a exposição televisiva, e que devem complementar as análises de construção dos modelos de exposição nos contextos de maior variabilidade dos dados de audimetria e de um fraco desempenho dos modelos de *ratings*. Devem ser identificadas as variáveis que podem apresentar uma maior capacidade predictiva do consumo televisivo.

Devem, também, ser equacionadas abordagens metodológicas alternativas quando se pretende o desenvolvimento dos modelos de *ratings*, uma vez que estas análises devem ser privilegiadas por incidirem nos dados que estão mais acessíveis, serem menos onerosas e mais rápidas. As análises univariadas de valores de *ratings* e de taxas médias de audiências para o total da TV podem, em determinados contextos, apresentar uma maior relação custo-benefício. O *Time Series Data Mining* poderá contribuir positivamente para o seu desenvolvimento, uma vez que incide na representação de ST, no *clustering* e na procura de sequências em grandes bases de dados.

Na perspectiva multivariada, foi considerado um conjunto muito abrangente de variáveis explicativas, quer para o desenvolvimento de modelos de *ratings* programáticos, quer para os dados comerciais. Recorrendo a informação que pode ser calculada a partir das bases de audimetria, foi possível construir modelos com uma elevada capacidade explicativa da evolução dos dados, mas que, do ponto de vista da previsão, não viabilizam resultados muito satisfatórios. Porém, os resultados que constam da bibliografia internacional não apresentam, igualmente, uma maior precisão (Danaher *et al.*, 2011; Danaher & Dagger, 2012; Nikolopoulos, 2010; Weber, 2002). Podem ser equacionados modelos alternativos, mas com uma maior ênfase ao nível das metodologias não lineares, como, por exemplo, as Árvores e as Regras de Decisão e as Redes Neurais, que são especialmente relevantes para a análise das observações atípicas.

CAPÍTULO VIII

REFERÊNCIAS BIBLIOGRÁFICAS

- Abdi, H. (2007). Bonferroni and Sidak corrections for multiple comparisons. In N. Salkind (Ed.), *Encyclopedia of measurement and statistics* (pp. 103-107). Thousand Oaks, CA: Sage.
- Agee, T. (1992). *The effects of television clutter and position on advertising recall*. Master Thesis, University of Auckland, Auckland, USA.
- Agostini, J. (1988). L'audience depuis vingt-cinq ans. Approche historique. In G. Santini & E. Santier, *L'audience et les medias* (pp. 39-67), Paris: IREP.
- Alavy, K., Gaskell, A., Leach, S., & Szymanski, S. (2010). On the edge of your seat: Demand for football on television and the uncertainty of outcome hypothesis. *International Journal of Sports Finance*, 5(2), 75-95.
- Almeida, J., Pinto, J., & Cruzeiro, E. (1973). A propósito do problema da causalidade em ciências sociais: O modelo de decomposição de proporções de Raymond Boudon. *Análise Social*, X(40), 734-777.
- Alwitt, L., & Prabhaker, P. (1992). Functional and belief dimensions of attitudes to television advertising: Implications for copytesting. *Journal of Advertising Research*, 32(5), 30-42.
- Amornchewin, R., & Kreesuradej, W. (2009). Mining dynamic databases using probability-based incremental association rule discovery algorithm. *Journal of Universal Computer Science*, 15(12), 2409-2428.
- Anacom (2010). *Serviço de Televisão por subscrição*. 3º Trimestre de 2010. Autoridade Nacional de Comunicações.
- Appice, A., Ceci, M., & Malerba, D. (2003). MR-SMOTI: A Data Mining system for Regression tasks tightly-coupled with a relational database. In *Proceedings of KDD-03, Washington DC, USA, 23-24 August 2003*, 17-27.

- Aris, A., & Bughin, J. (2005). *Managing media companies. Harnessing Creative Value*. London: John Wiley and Sons, Ltd.
- Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, 4, 40-79.
- Atkeson, G., Moore, W., & Schaal, S. (1997). Locally weighted learning. In D. W. Aha (Ed.), Special issue on lazy learning (edited collection). *Artificial Intelligence Review*, 11(1-5), 11-73.
- Atramentov, A. (2003). Multi-relational decision tree algorithm implementation and experiments. Master Thesis in Computer Science, Iowa State University, Iowa, USA.
- Atramentov, A., Leiva, H., & Honavar, V. (2003). A multi-relational decision tree learning algorithm – Implementation and experiments. *Inductive Logic Programming. Lecture Notes in Computer Science*, 2835, 38-56.
- Barwise, T. (1986). Repeat-viewing of prime-time series. *Journal of Advertising Research*, 26(4), 9-14.
- Barwise, T., Ehrenberg, A., & Goodhardt, G. (1982). Glued to the box?: Patterns of repeat viewing. *Journal of Communication*, 32(4), 22-29.
- Barwise, P., & Ehrenberg, A. (1988). *Television and its audience*. London: Sage Publications.
- Basheer, I. & Hajmeer, M. (2000). Artificial neural networks: Fundamentals, computing, design and application, *Journal of Microbiological Methods*, 43(1), 3–31.
- Becker, B., Kohavi, R., & Sommerfield, D. (2001). Visualizing the simple bayesian classifier. In U. Fayyad, G. Grinstein & A. Wierse (eds). *Information Visualization in Data Mining and Knowledge Discovery*. San Francisco: Morgan Kaufmann Publishers.
- Benaroya, G. (1995). Director CBC. Express: Um sistema de previsão de la audiência TV. In M. Zarate (1997), *Los modelos de predicción de las audiencias de los programas*. Tesis Doctoral. Madrid: Universidad Complutense de Madrid.

- Bensi, R. (2013). *Conceito de classe e estratificação social por Anthony Giddens*. Disponível em <http://www.artigonal.com/ensino-superior-artigos/conceito-de-classe-e-estratificacao-social-por-anthony-giddens-1619341.html>
- Berkhin, P. (2006). A survey of clustering data mining techniques. In J. Kogan, C. Nicholas & M. Teboulle (Eds.), *Grouping multidimensional data: Recent advances in clustering* (pp. 25-71). Berlin / Heidelberg: Springer.
- Beville, H. (1988). *Audience ratings: Radio, television, cable*. (revised edition). Hillsdale, NJ: Lawrence Erlbaum.
- Billard, L., & Diday, E. (2000). Regression analysis for interval-valued data. In H. Kiers, J. Rasson, P. Groenen & M. Schader (Eds.), *Data analysis, classification and related methods* (pp. 369-374). Berlin / Heidelberg: Springer.
- Blankenship, A. (1977). *Professional telephone surveys*. New York: McGraw-Hill Book Company.
- Bock, H.-H., & Diday, E. (Eds.) (2000). *Analysis of symbolic data: Exploratory methods for extracting statistical information from complex data*, (1st ed). Berlin: Springer-Verlag.
- Boemer, M. (1987). Correlating lead-in show ratings with local television news ratings. *Journal of Broadcasting & Electronic Media*, 31(1), 89-94.
- Borg, I. & Groenen, P. (2005). *Modern Multidimensional Scaling: Theory and Applications*, 2nd ed. New York: Springer.
- Bortz, J. (1986). Zur messung der attraktivität von fernsehsendungen. Regression sresiduen als Indikatoren. *Rundfunk und Fernsehen*, 34, 64-72.
- Branco, J. (2010). Estatísticos e mineiros (de dados): Inseparáveis de costas voltadas?. *Boletim Sociedade Portuguesa de Estatística, Primavera*, 40-43.
- Breiman, L., Friedman, J., Olshen, R., & Stone, J. (1984). *Classification and Regression trees*. Belmont, California, USA: Wadsworth Int. Group.
- Brosius, H., & Wober, M. (1992). The loyalty of television viewing: How consistent is TV viewing behaviour? *Journal of Broadcasting & Electronic Media*, 36(3), 321-335.

- Bryant, J., & Zilmann, D. (1985). *Perspectives on media effects*. Hillsdale, NJ: Lawrence Erlbaum.
- Bukharin, N. (1933). *A teoria do materialismo histórico. Manual popular de sociologia marxista*. Disponível em <http://www.marxists.org/portugues/bukharin/1921/teoria/index.htm>
- Bulck, J. (2006). Television news avoidance: Exploratory results from a one-year follow-up study. *Journal of Broadcasting and Electronic Media*, 50(2), 231-252.
- Bunge, M. (1981). *Teoria y realidad*. Barcelona: Editorial Ariel.
- Bunge, M. (1982). *Filosofía de la física*. Barcelona: Editorial Ariel.
- Bunge, M. (1989). *La investigación científica: Su estrategia y su filosofía*. Barcelona: Ediciones Ariel.
- Bunge, M. (2005). *La ciencia: Su método y su filosofía*. Buenos Aires: Editorial Sudamericana.
- Bunn, D. (1982). Audience presence during breaks in TV programs. *Journal of Advertising Research*, 22(5), 35-39.
- Buraimo, T. (2007). *An econometric analysis of market size and audience demand in football*. PH,D Thesis, Lancaster University Management School (LUMS), Lancaster, United Kingdom.
- Buraimo, B., & Simmons, R. (2009). A tale of two audiences: Spectators, television viewers and outcome uncertainty in Spanish football. *Journal of Economics and Business*, 61(4), 326-338.
- Burnett, J. (1991). Examining the media habits of the affluent elderly. *Journal of Advertising Research*, 31(5), 33-41.
- Bush, A., & Leigh, J. (1984). Advertising on cable versus traditional networks. *Journal of Advertising Research*, 24(2), 33-38.
- Cádima, F. (1996). *O fenómeno televisivo*. Lisboa: Circulo de Leitores.
- Cádima, F. (1998). *Portugal e os desafios da comunicação. Em Portugal na transição do milénio – Colóquio Internacional*. Lisboa: Fim de Século Edições.

- Caldeira (2000). Critical realism: A philosophical perspective. *Episteme*, Ano II(5-6), 73-78.
- Callan, J. (2009). *Global ad spending to drop most in 29 years, hurting TV*, *Print*. Disponível em <http://www.bloomberg.com/apps/news?pid=newsarchive\&sid=anWJD4QmrsE>
- Callejo, J. (1995). *La audiência activa. El consumo televisivo: Discursos y estrategias*. Madrid: Centro de Investigación Sociológica (CIS): Siglo vientiuno de España.
- Cameron, A., & Pravin, K. (2005). *Microeconometrics: Methods and applications*. New York: Cambridge University Press: Sections 4.1-4.4.
- Cannon, H., & Riordan, E. (1994). Effective reach and frequency: Does it really make sense?. *Journal of Advertising Research*, 34(2), 19-28.
- Cannon, H., & Leckenby, J., & Abernethy, A. (2002). Beyond effective reach frequency: Evaluating media schedule using frequency value planning. *Journal of Advertising Research*, 42(6), 33-47.
- Carnap, R. (1950). *Logical foundations of probability*. The University of Chicago Press Corp.
- Cassetti, F., & Chio, F. (1997). *Análisis de la televisión: Instrumentos, metodos y practicas de investigación*. Barcelona: Paidós.
- Cattin, P., Festa, R., & Le Diberder, A. (1994). A model for forecasting the audience of TV Programs. In ARF (Advertising Research Foundation) & ESOMAR (European Society for Opinion and Marketing Research) (Eds.), *Symposium on Worldwide Electronic and Broadcasting Audience Research*. Vol I (pp. 513-524). Paris.
- Cendrowska, J. (1987). PRISM: An algorithm for inducing modular rules. *International Journal of Man-Machine Studies*, 27(4), 349-370.
- Chakrapani, C. (1993). The relevance of buyer behaviour models. *Imprints, Magazine of the PMRS*. April. Disponível em <http://www.chuckchakrapani.com/Articles/>
- Chambers, R., & Skinner, C. (2003). *Analysis of Survey Data*. New York: John Wiley & Sons.

- Chaniac, R. (1988). La mesure d'audience de la télévision. *Dossiers de l'Audiovisuel*, 22, 37-39.
- Chouakria, A., Cazes, P., & Diday, E. (2000). Symbolic Principal Components Analysis. In H.-H Bock, & E. Diday (Eds.), *Analysis of symbolic data* (200–212). Berlin: Springer.
- Chow, T. & Cho, S: (2007). *Neural networks and computing*. London: Imperial College Press.
- Clancy, M. (1994). The television audience examined. *Journal of Advertising Research*, 34(4), 76-87.
- Clancy, K., & Kweskin, D. (1971). TV commercial recall correlates. *Journal of Advertising Research*, 11(2), 18-20.
- Cleveland, W. & Loader, C. (1996). Smoothing by local Regression: Principles and methods. In W. Härdle & M. Schimek (Eds.), *Statistical theory and computational aspects of smoothing*, (10-49). New York: Springer.
- Cobb, C. (1985). *Television clutter and advertising effectiveness*. Comunicação apresentada na *American Marketing Association Educator's Conference*. Chicago: American Marketing Association.
- Comstock, G. (1980). *Television in America*. Beverly Hills, London: Sage.
- Cooper, R. (1993). An expanded, integrated model for determining audience exposure to television. *Journal of Broadcasting and Electronic Media*, 37(4), 401-418.
- Cooper, R. (1996). The status and future of audience duplication research: An assessment of ratings-based theories of audience behaviour. *Journal of Broadcasting and Electronic Media*, 40(4), 96-111.
- Cooper, R., & Tang, T (2009). Predicting audience exposure to television in today's media environment: An empirical integration of active-audience and structural theories. *Journal of Broadcasting & Electronic Media*, 53(3), 400-418.
- Corder, G., & Foreman, D. (2009). *Nonparametric statistics for non-statisticians: A step-by-step approach*. New York: Wiley.

- Craig, C., & Ghosh, A. (1993). Using household level viewing data to maximize effective reach. *Journal of Advertising Research*, 33(1), 38-47.
- CRECS (2006). Deliberação 1-l/2006. Conselho Regulador da Entidade Reguladora para a Comunicação Social.
- Cunningham, P., & Delaney, S. (2007). *K-Nearest Neighbor Classifiers*. Technical Report UCD-CSI-2007-4, School of Computer Science and Informatics. University College Dublin, Ireland.
- Danaher, P. (1995). What happens to television ratings during commercial break. *Journal of Advertising Research*, 35(1), 37-47.
- Danaher, J., & Beed, T. (1993). A coincidental survey of people meter panelists: Comparing what people say with what they do. *Journal of Advertising Research*, 33(3), 86-92.
- Danaher, P., & Dagger, T. (2012). Using a nested logit model to forecast television ratings. *International Journal of Forecasting*, 28(3), 607-622.
- Danaher, P., Dagger, T., & Smith, M. (2011). Forecasting TV Ratings. *International Journal of Forecasting*, 27(4), 1215-1240.
- Danaher, P., & Mawhinney, D. (2001). Optimizing television program schedules using choice modeling. *Journal of Marketing Research*, 38(3), 298-312.
- Darmon, R. (1976). Determinants of TV viewing. *Journal of Advertising Research*, 16(6), 17-24.
- Das, G., Mannila, H. & Smyth, P. (1998). Rule discovery from time series. In *Proceedings of the Conference on Knowledge Discovery and Data Mining (KDD-98)*, New York City, New York, EUA, 27-31 August 1998 (16-22).
- Datia, N. (2002). *Ferramenta de aprendizagem indutiva: Previsão de audiências televisivas com Árvores de Regressão e de classificação*. Relatório do projecto final do Curso de Engenharia de Informática e de Computadores, ISEL-DEETC-SSIC, Lisboa, Portugal.
- Davis, D., & Walker, J. (1990). Countering the new media: The resurgence of share maintenance in primetime network television. *Journal of Broadcasting & Electronic Media*, 34(4), 487-493.

- De Ville, B. (2006). *Decision trees for business intelligence and data mining: Using SAS Enterprise Miner*. Cary, N. C: SAS Institute.
- Diday, E., & Noirhomme-Fraiture, M. (2008). *Symbolic data analysis and the SODAS software*. Chichester: John Wiley & Sons.
- Dilão, R. (1995). Em ciência dos sistemas complexos. *Técnica*, 1, 5-18.
- Di Domizio, M. (2010). *Competitive balance and TV audience: An empirical analysis of the Italian Serie A*. Working paper 2010/64. Teramo, Italy: University of Teramo.
- Dickson, P. (1991). Grp: A case of mistaken identity. *Journal of Advertising Research*, 31(1), 55-59.
- Ding, C., & Xiaofeng, H. (2004). K-means clustering via principal component analysis. In *Proceedings of Int'l Conf. Machine Learning (ICML- 04)*, Banff, Alberta, Canada, 4-8 July 2004 (225-232).
- Dominick, J., Sherman, L., & Copeland, A. (1990). *Broadcasting cable and beyond. An introduction to modern electronic media*. EE.UU: McGraw-Hill.
- Dougherty, G. (2013). *Pattern recognition and classification: An introduction*. New York, NY: Imprint: Springer.
- Duda, R., Hart, P., & Stork, D. (2000). *Pattern Classification* (2nd ed.). Wiley Interscience. ISBN 0-471-05669-3. MR 1802993.
- Durand, J. (1982). *La mesure automatique de l'audience de la télévision*. Apresentado: Seminario internacional sobre investigação de audiências de jornais, radio e televisão, organizado pela universidade de Navarra em Pamplona em Junho 1982.
- Durkheim, E. (1980). *As Regras do método sociológico*. Lisboa: Editorial Presença.
- Dzeroski, S. (2005). Relational data mining. In O. Mainon & L. Rokach (Eds.), *Data mining and knowledge discovery handbook* (pp. 869-898). Tel-Aviv, Israel: Springer.
- Dzeroski, S., & Lavrac, N. (2001). *Relational Data Mining*. Berlin: Springer.

- Easterby-Smith, M. et al. (1999). *A pesquisa gerencial em administração: Um guia para monografias, dissertações, pesquisas internas e trabalhos em consultoria*. São Paulo: Pioneira.
- Eastman, S., & Newton, G. (1996). Promoting prime time programs in mega sporting events. *Journal of Broadcasting & Electronic Media*, 40(3), 366-388.
- Eastman, S., & Newton, G., & Gregory, D. (1998). The impact of structural salience within on air promotion. *Journal of Broadcasting & Electronic Media*, 42(1), 50-79.
- Eastman, S., Newton, G., Riggs, K., & Neal-Lunsford, J. (1997). Accelerating the flow: A transition effect in programming theory?. *Journal of Broadcasting & Electronic Media*, 41(2), 265-283.
- Eastman, S., & Otteson, J. (1994). Promotion increases ratings, doesn't It? The impact of program promotion in the 1992 Olympics. *Journal of Broadcasting & Electronic Media*, 38(3), 307-322.
- Ehrenberg, A., & Wakshlag, J. (1987). Repeat-viewing with people meters. *Journal of Advertising Research*, 27(1), 9-13.
- ERC (2010). Relatório de Regulação 2009. Entidade Reguladora para a Comunicação Social.
- Esling, P., & Agon, C. (2012). Time-series data mining. *ACM Computing Surveys*, 45(1), 1-34.
- Espinosa, M. (2006). *La planificación de medios de comunicación de massas*. Madrid: Mc Graw Hill.
- Esposito, F., Malerba, D., & Tamma, V. (2000). Dissimilarity Measures for Symbolic Objects. In H. Bock & E. Diday (eds.): *Analysis of Symbolic Data. Exploratory Methods for extracting Statistical Information from Complex Data, Series: Studies in Classification, Data Analysis, and Knowledge Organisation*, Vol. 15, 165-185. Berlin: Springer-Verlag.
- Farris, P., & Parry, M. (1991). Clarifying some ambiguities regarding grp and average frequency. A comment on "GRP: A case of mistaken identity". *Journal of Advertising Research*, 31(1), 75-77.
- Faustino (2007). Becoming a broadcasting leader in 10 years. *International Journal on Media Management*, 9(4), 151-163.

- Ferguson, D. (1992). Channel repertoire in the presence of remote control devices, VCRs, and cable television. *Journal of Broadcasting & Electronic Media*, 36(1), 83-91.
- Ferguson, D., & Perse, E. (1993). Media and audience influences on channel repertoire. *Journal of Broadcasting & Electronic Media*, 37(1), 31-47.
- Fernandes, A. (2001). *Televisão do público*. Coimbra: Minerva.
- Fernandes, A. (2008). A televisão na sociedade em rede. Comunicação apresentada no VI Congresso Português de Sociologia – *Mundos sociais: Saberes e práticas*. Universidade Nova, Lisboa: Portugal, 25-28 de Junho.
- Ferreira, M., Reis, N., & Santos, J. (2011). TVI: O turnaround até à liderança das audiências. Caso de Estudo nº 7, globADVANTAGE. Center of Research in International Business & Strategy.
- Filho, J. (2008). Causalidade, indeterminismo e imprevisibilidade. Porque o conceito de causa é tão importante?. *Revista Brasileira de Ensino da Física*, 30(3), 3304.
- Filho, J., & Selleri, F. (2006). Popper's Propensities in Quantum Physics. In M. Alai, & G. Tarozzi (Eds.), *Proceedings of the Conference Karl Popper Philosopher of Science* (pp. 277-295). Cesena, 27-30 de Outubro de 1994. Soveria Mannelli: Rubbettino.
- Finetti, B. (1937). Foresight: its logical laws, its subjective sources. In H. E. Kyburg & H. E. Smokler (Eds.), *Annales de l'Institut Henri Poincaré* (traduzido em 1964, in *Studies in Subjective Probability*). New York: Wiley.
- Fitzsimons, M., Khabaza, T., & Shearer, C. (1993). The application of rule induction and neural networks for television audience prediction. In *Proceedings of Esomar/Emac/ Afm Symposium on Information based decision making in marketing, Paris, France, November 1993* (pp. 69-82).
- Fonseca, J. (2008). Os métodos quantitativos na sociologia: Dificuldades de uma metodologia de investigação. Comunicação apresentada no VI Congresso Português de Sociologia – *Mundos sociais: Saberes e práticas*. Universidade Nova, Lisboa: Portugal, 25-28 de Junho.
- Ford, J., Pitt, L., & Okazaki, S. (2008). The art of reviewing. *International Journal of Advertising*, 27(2), 331-340.

- Forrest, D., Simmons, R., & Buraimo, B. (2005). Outcome uncertainty and the couch potato audience. *Scottish Journal of Political Economy*, 52(4), 641-661.
- Fox, J. (2000). *Multiple and generalized nonparametric Regression*. Thousand Oaks, CA: Sage.
- Foxall, G. (2009). Ascribing intentionality. *Behavior and Philosophy*, 37, 217-222.
- Gantz, W. (1985). Exploring the role of married life. *Journal of Broadcasting & Electronic Media*, 29(1), 65-78.
- Garland, I. (1996). On the possibility of predicting ratings reliably. In: ARF (Advertising Research Foundation) & ESOMAR (European Society for Opinion and Marketing Research) (Eds.), *Symposium on Worldwide Electronic and Broadcasting Audience Research* (pp. 233-247). San Francisco.
- Geel, P., Hoogstrate, H., & Foley, T. (2003). TV ratings: Can predictions beat reality. *Admap*, 436, 38-39.
- Gensch, D., & Shaman, P. (1980). Models of competitive television ratings. *Journal of Marketing Research*, 17(3), 307-315.
- Gensch, D., & Shaman, P. (1980). Predicting TV ratings. *Journal of Advertising Research*, 20(4), 85-92.
- Gerhards, M., & Klingler, W. (2005). Programmangebote und spartennutzung im fernsehen: ergebnisse 2004 und ein fünfjahresvergleich auf basis der AGF/GfK-Programmcodierung. *Media Perspektiven*, 11, 558-569.
- Getoor, L. (2005). Link mining: A survey. *ACM SIGKDD Explorations Newsletter*, 7(2), 3-12.
- Getoor, L., & Taskar, B. (2007). *Introduction to Statistical Relational Learning* (Adaptive Computation and Machine Learning series). The MIT Press. Cambridge, MA.
- Ghosh, A. (2008). Where television audience measurement needs to go in China. *Admap (China supplement)*, February, 20-22.

- Gill, J. (2000). Managing the capture of individuals viewing within a peplemeter service.(research practices that need to be implemented on a TV viewing peplemeter panel). *International Journal of Market Research. World Advertising Research Center Ltd.* Retrieved November 27, 2013 from HighBeam Research: <http://www.highbeam.com/doc/1G1-67543717.html>
- Gil, A. (1999). *Métodos e técnicas de pesquisa social*. São Paulo: Atlas.
- Givon, M., & Grosfeld-Nir, A. (2008). Using partially observed Markov processes to select optimal termination time of TV shows. *Omega*, 36(3), 477–485.
- Godoy, A. (1995). Introdução à pesquisa qualitativa e suas possibilidades. *Revista de Administração de Empresas*, 5(2), 57-63
- Goetz, P., & Lecompte, M. (1988). *Etnografia y diseño qualitativo en la investigación educativa*. Madrid: Morata.
- Goldberg. M. et al. (1993). Music, music videos and wearout. *Psychology and Marketing*, 10(1), 1-13.
- Goodhart, G. (1966). The constant in duplicated television viewing between and within channels. *Nature*, 212, 1616.
- Goodhart, G., Ehrenberg, A., & Collins, M. (1975). *The television viewing audience: Pattern of viewing*. Lexington, MA: Lexington Books.
- Goodhart, G., Ehrenberg, A., & Collins, M. (1987). *The television audience: Patterns of viewing* (2nd ed.). Westmead, UK: Gower.
- Gordon, L. & Patzer, G. (1991). Multiple dimensions of performance for 30 and 15 second commercials. *Journal of Advertising Research*, 31(4), 18-25.
- Gowda, K. & Ravi, T. (1995). Divisive clustering of symbolic objects using the concepts of both similarity and dissimilarity. *Pattern Recognition*, 28(8), 1277-1282.
- Green, A. (2007). *Understanding the TV viewer*. WARC Best Practice, November. Disponível em http://www.ipsos.com/mediact/sites/ipsos.com.mediact/files/pdf/warc_best_practice-understanding_the_tv_viewer-nov_2007.pdf.

- Greenacre, M. (2007). *Correspondence analysis in practice*. (2nd edition). Boca Raton, FL: Chapman & Hall/CRC.
- Greenacre, M. & Blasius, J. (2006). *Multiple correspondence analysis and related methods*. London: Chapman & Hall/CRC Press.
- Grenberg, B., Sherry, J., Busselle, R., Hnilo, L., & Smith, S. (1997). Daytime television talk shows: Guests, content and interactions. *Journal of Broadcasting & Electronic Media*, 41(3), 426-440.
- Groenen, P., Winsberg, S., Rodriguez, O., & Diday, E. (2005). *Multidimensional scaling of interval dissimilarities*. *Econometric Report*, 2005-15. Rotterdam: Erasmus University.
- Groenen, P., Winsberg, S., Rodriguez, O., & Diday, E. (2006). I-Scal: Multidimensional scaling of interval dissimilarities. *Computational Statistics and Data Analysis*, 51, 360-378.
- Grove, W. Zald, D., Lebow, B., Snitz, B., & Nelson, C. (2000). Clinical versus mechanical predictions. A meta-analysis. *Psicological Assessment*, 12(1), 19-30.
- Guimarães, G. (2000). The induction of temporal grammatical rules from multivariate time series. In A. L. Oliveira (Ed.), *Grammatical inference: Algorithms and applications. Proceedings of the 5th International Colloquium, ICGI 2000* (pp. 127-140). Lisboa, Portugal, 11-13 de Setembro.
- Das, G., Gunopulos, D. & Mannila, H. (1997). Finding Similar Time Series. In *Proceedings of First European Symposium (PKDD '97), Trondheim, Norway, June 24-27, 1997* (pp.88-100). Berlin / Heidelberg: Springer.
- Hair, J., Anderson, R., Tatham, R. & Black, W. (1998). *Multivariate Data Analysis*. New Jersey: Prentice-Hall.
- Hammer, P., Riebe, E., & Kennedy, R. (2009). How clutter affects advertising effectiveness. *Journal of Advertising Research*, 49(2), 159-163.
- Han, J., & Kamber, M. (2001). *Data Mining: concepts and techniques*. San Francisco: Morgan-Kaufmann Academic Press.
- Hand, D. (1998). Data mining: statistics and more? *The American Statistician*, 52, 112-118.

- Hall, M., Eibe, F., Holmes, G. Pfahringer, B., Reutemann, P. & Witten, I. (2009). The WEKA data mining software: An update. *SIGKDD Explorations*, Volume 11, Issue 1.
- Hand, D. (1999). Statistics and data mining: Intersecting disciplines. *SIGKDD Explorations*, 1, 16-19.
- Hardle, W. (1990). *Applied nonparametric regression*. Cambridge: University Press.
- Hartley, J. (2000). *Los usos de la televisión*. Barcelona: Paidós.
- Harwood, J. (1997). Viewing age: Lifespan identity and television viewing choices. *Journal of Broadcasting & Electronic Media*, 41(2), 203-213.
- Hauser, J. (1978). Consumer preference axioms: behavioural postulates for describing and predicting stochastic choice. *Management Science*, 24(13), 1331-1341.
- Hastie, T., & Loader, C. (1993). Local regression: Automatic kernel carpentry. *Statistical Science*, 8(2), 120-143.
- Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The elements of statistical learning*. New York: Springer Verlag.
- Headen, K., Klompmaker, J., Teel, J., & Jessie, G. (1977). Predicting audience exposure to spot TV advertising schedules. *Journal of Marketing Research*, 14(1), 1-9.
- Headen, R., Klompmaker, J., & Rust, R. (1979). The duplication of viewing law and television media schedule evaluation. *Journal of Marketing Research*, 16(3), 333-340.
- Heeter, C. (1985). Program choice with abundance of choice: A process model. *Human Communication Research*, 12(1), 126-152.
- Heeter, C. (1988). *The choice process model*. In C. Heeter & B. Greenberg (Eds.), *Cable viewing* (pp.11-32). Norwood, Nj: Ablex.
- Heeter, C., & Greenberg, B. (1985). Profiling the zappers. *Journal of Advertising Research*, 25(2), 15-19.

- Heller, A. (1989). De la hermeneutica en las ciencias sociales a la hermeneutica de las ciencias naturales. In A. Heller, & F. Fehér, F. (Eds.), *Políticas de posmodernidade. Ensayos de la crítica cultural* (pp. 52-98). Barcelona: Península.
- Henriksen, F. (1985). A new model of the duplication of television viewing: A behaviourist approach. *Journal of Broadcasting & Electronic Media*, 29(2), 135-145.
- Henry, M., & Rinne, H. (1984). Predicting program shares in new time slots. *Journal of Advertising Research*, 24(2), 9-17.
- Henry, H., & Rinne, H. (1984). Offensive versus defensive TV programming Strategies. *Journal of Advertising Research*, 24(3), 45-56.
- Herbeg, J. (1997) *La evolución del pensamiento mítico al pensamiento racional*. Barcelona: Herder.
- Hinkelmann, K., & Kempthorne, O. (2008). *Design and analysis of experiments. Volume I: Introduction to Experimental Design* (Second Ed.). Hoboken, New Jersey: John Wiley & Sons. ISBN 978-0-471-72756-9.
- Huertas, A. (1997). *Como se miden las audiências en televisión*. Barcelona: CIMS.
- Huertas, A. (2002). *La audiência investigada*. Barcelona: Gedisa editorial.
- Hyman, H. (1984). *Diseño y análisis de las encuestas sociales*. Buenos Aires: Amorrortu.
- Igartua, J., & Badillo, A. (eds) (2003). *Audiência y medios de comunicación*. Salamanca: Universidad de Salamanca.
- Jaeyik, K. (2012). Dissimilarity measures and divisive clustering for symbolic multimodal-valued data. *Computational Statistics & Data Analysis*, 56(9), 2795-2808.
- Jain, A. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8), 651-666.
- Jauset, J. (2008). *La audiencia en la televisión*. Barcelona: UOC.
- Jensen, F. (2001). *Bayesian networks and decision graphs*. New York: Springer-Verlag.

- Johnson, R., & Cobb-Walgren, C. (1994). Aging and the problem of TV clutter. *Journal of Advertising Research*, 34(4), 54-62.
- Johnsen, H., & Solvoll, M. (2007). The demand for televised football. *European Sport Management Quarterly*, 7(4), 311-335.
- Johnson, R., & Cobb-Walgren, C. (1994). Aging and the problem of TV clutter. *Journal of Advertising Research*, 34(4), 54-62.
- G-Feng, G., Jing, L., & Wei-Feng, B. (2007). An efficient relational decision tree classification algorithm. In *Proceedings of the Third International Conference on Natural Computation, Haikou, China, 24-27 August 2007* (pp.530-534).
- Kamerer, D. (1987). *Factors mediating repeat viewing of television programs*. Comunicação apresentada no encontro da Association for Education in Journalism and Mass Communication. San Antonio, TX, Agosto.
- Kaplan, B. (1985). Zapping - The real issue is communication. *Journal of Advertising Research*, 25(2), 9-12.
- Katkovnik, V. (1995). Nonparametric local polynomial approximation of the time-varying frequency and amplitude. *Communications in Statistics - Theory and Methods*, 24(12), 3001-3025.
- Kaufman, L., & Rousseeuw, P. J. (2008). References. In *Finding groups in data: An introduction to cluster analysis*. Hoboken, NJ, USA: John Wiley & Sons, Inc. doi: 10.1002/9780470316801.refs.
- Kass, G. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics* 29(2), 119-127.
- Kelton, C., & Schneider-Stone, L (1998). Optimal television schedules in alternative competitive environments. *European Journal of Operational Research*, 104, 451-473.
- Kent, R. (1995). Competitive versus non competitive clutter in TV advertising. *Journal of Advertising Research*, 35(2), 49-56.
- Ketterlin, A. (1997). Clustering sequences of complex objects, In *Proceedings of the 3rd Int. Conf. on Knowledge Discovery and Data Mining (KDD-97)*, Newport Beach, California, USA, 14-17 August 1997 (pp. 215 -218).

- Keppel, G., & T. Wickens. (2004). *Design and analysis. A researcher`s handbook* (4th edition). Upper Saddle River, NJ: Prentice Hall.
- Kibler, D., Aha, D., & Albert, M. (1989). Instance-based prediction of real-valued attributes. *Computational Intelligence*, 5, 51-57.
- Kim, C. (2002). Identifying viewer segments for television programs, *Journal of Advertising Research*, 42(1), 51-66.
- Kim, J. & Billard, L. (2012) Dissimilarity measures and divisive clustering for symbolic multimodal-valued data. *Computational Statistics & Data Analysis*, 56, 2795-2808.
- Kirsh, D., & Banks, S. (1962). Program types defined by factor analysis. *Journal of Advertising Research*, 2(3), 29-31.
- Kleinbaum, D. (2010). *Logistic Regression*. New York: Springer-Verlag.
- Kleinbaum, D., Kupper, L., Muller, K., & Nizam, A. (1998). *Applied regression analysis and other multivariable methods*. Pacific Grove: Duxbury Press.
- Knobbe, A., Siebes, A., & Van der Wallen, D. (1999). Multi-Relational decision tree induction. In J. M. Zytkow, & J. Rauch (Eds.), *Principles of data mining and knowledge discovery. Proceedings of the Third European Conference (PKDD'99)*, (pp. 378-383). Prague, Czech Republic, September 15-18. Berlin: Springer.
- Kohonen, T. (2001). *Self-organizing maps*. (3rd Extended Ed). New York: Springer-Verlag.
- Kramer, S. & Widmer, G. (2001). Inducing Classification and Regression Trees in First-Order Logic. In: S. Dzeroski & N. Lavrac (eds.). *Relational Data Mining: Inductive Logic Programming for Knowledge Discovery in Databases* (149-159). Berlin/Heidelberg/New York/Tokyo: Springer Verlag.
- Krugman. E. (1972). Why three exposures may be enough. *Journal of Advertising Research*, 12(6), 11-14.
- Kulkarni, V., & Kekre, H. (2010). Speaker identification by using vector. Quantization. *International Journal of Engineering Science and Technology*, 2(5), 1325-1331.
- Landé, A. (1968). *Nuevos fundamentos de la mecánica cuantica*. Madrid: Editorial Tecnos.

- Langley, P., Iba, W., & Thompson, W. (1992). An analysis of bayesian classifier. In *Proceeding of the 10th National Conference on Artificial Intelligence, San Jose, California, 12-16 July 1992* (pp.223-228).
- Lauro, N., & Palumbo, F. (2000). Principal component analysis of interval data: A symbolic data analysis approach. *Computation Statistics*, 15(1), 73-87.
- Laxman, S., & Sastry, P. (2006). A survey of temporal data mining. *Sadhana*, 31(2), 173-198.
- Lehman, D. (1971). Television show preference: Application of a choice model. *Journal of Marketing Research*, 8(1), 47-55.
- Leiva, H., Atramentov, A., & Honavar, V. (2002). Experiments with MRDTL – A multi-relational decision tree learning algorithm. Technical Report ISU-CS-TR 02-12, Computer Science- Iowa: Iowa State University.
- Lemire, D. (2006). *A better alternative to piecewise linear time series segmentation*. Technical Report. Disponível em <http://arxiv.org/pdf/cs/0605103.pdf>
- Leon, J. (1990). *A medição da audiência no jornalismo audiovisual antes do ano 2000*. Bilbao: Serviço Editorial Universidade do País Vasco.
- Lee, S., & Lumpkin, J. (1995). Differences in attitudes toward TV advertising: VCR usage as a moderator. *International Journal of Advertising*, 11(4), 333-42.
- Li, Q., & Racine, J. (2007). *Nonparametric econometrics: Theory and practice*. Princeton: Princeton University Press. ISBN 0691121613.
- Lipovetsky, G. (2000). *A era do vazio. Ensaios sobre o individualismo contemporâneo*. (Duodécima Edição). Barcelona: Anagrama.
- Litman, B., & Kohl, L. (1992). Network rerun viewing in the age of new programming services. *Journalism Quarterly*, 69(2), 383-391.
- Lima, V. (2007). O lugar da psicologia na sociologia de Durkheim. *Qualit@s Revista Electrónica*, 6(1). Disponível em <http://revista.uepb.edu.br/index.php/qualitas/article/view/74/67>
- Lochte, R., & Warren, J. (1989). A channel repertoire for TV satellite viewers. *Journal of Broadcasting & Electronic Media*, 33(1), 91-95.

- Longman, K. (1997). If not effective frequency, then what? *Journal of Advertising Research*, 37(4), 44-51.
- McDonald, C. (1995). What do we know about...? The context of the commercial. Programme involvement, and how it affects response to TV advertisements. *Admap*, 30, 16.
- Mackensie, S., Lutz, R., & Belch, G. (1986). The role of attitude toward the ad as a mediator of advertising effectiveness: A test of competing Explanations. *Journal of Marketing research*, 23(2), 130-43.
- Madinaveitia, E. (1989). El índice de atracción. In AEDEMO, 46º Seminario AEDEMO. 5º Seminario sobre Audiencia de Televisión (pp. 99-125). San Sebastián, 2 e 3 de fevereiro. Barcelona: AEDEMO.
- Magidson, J. (1993). The use of the new ordinal algorithm in CHAID to target profitable segments. *The Journal of Database Marketing*, 1, 29-44.
- Magidson, J. (1994). The CHAID approach to segmentation modeling: Chi-squared automatic interaction detection. In R. Bagozzi (Ed.), *Advanced Methods of Marketing Research* (pp. 118-159). Cambridge (Massachusetts): Blackwell Business.
- Maia, A., Carvalho, F., & Ludemir, T. (2006). Symbolic interval time series forecasting using a hybrid model. In Proceedings of SBRN '06, Ninth Brazilian Symposium on Neural Networks, Ribeirao Preto, Brazil, 23-27 October 2006 (pp. 202-207).
- Makienko, I. (2012). Effective frequency estimates in local media planning practice. *Journal of Targeting, Measurements and Analysis in Marketing*, 20(1), 57-65.
- Makridakis, S., Wheelwright, S., & Hyndman, R. (1995). *Forecasting: Methods and applications*. New York: John Wiley & Sons.
- Malara, C., Armellini, A., & Monotti, S. (2004). ESOMAR. Can audiences of new or existing programs be predicted reliably? Implementing a forecasting model to predict program ratings and campaign schedules reach and frequency. Television Audience Conference, Geneva, June.
- Mancisidor, D. (1984). *La empresa de radio en USA*. Pamplona: Universidad de Navarra.

- Manyntz, R., Holm, K., & Hubner, P. (2005). *Introducción a los métodos de la sociologia empirica*. Madrid: Allianza.
- Mares, M., Oliver, M., & Cantor, J. (2008). Age differences in adult's emotional motivation for exposure to films. *Media Psychology*, 11, 488-511.
- Mariet, F. (1990). Le marché télévisuel et la démocratie de l'audience. In *Communications* (pp. 57-69). Paris: Le Centre d'Etudes Transdisciplinaires, EHESS, nº 51.
- Marktest (2009). *Uma década de audiências em Portugal*. Marktest Consulting. Disponível em <http://www.marktest.com/wap/a/p/id~fb.aspx>
- Martin, A. (1989). *O passado, hoje e amanhã dos Audímetros*. Em AEDEMO, 46º Seminário, 5º Seminário sobre as audiências de televisão (pp. 73-99). San Sebastián, 2 e 3 de fevereiro. Barcelona: AEDEMO.
- McDowell, W., & Dick, J. (2003). Has lead-in lost its punch? An analysis of prime time inheritance effects: Comparing 1992 and 2002. *International Journal of Media Management*, 5, 285-293.
- McMellon, C. & Eftimov, I. (2004). Exploring the determinants of the media negotiation process: A qualitative study. *Journal of Current Issues and Research in Advertising*, 26(2), 27-39.
- McQuail, D. (2000). *Introducción a la teoria de la comunicación de massas*. Barcelona: Paidós.
- Meehl, P. (1954). *Clinical versus statistical prediction*. Minneapolis: University of Minnesota Press.
- Merriam, S. (1998). *Qualitative research and case study applications in education*. São Francisco, CA: Jossey-Bass.
- Meyer, D., & Hyndman, R. (2005). The accuracy of television network rating forecasts: The effects of data aggregation and alternative models. *Model Assisted Statistics and Applications*, 1(3), 147-155.
- Meyer, D., & Muthaly, S. (2008). New measures and a new model for television network loyalty (MOTNL). *Marketing Bulletin*, 19(1), 1-19.

- Mises, R. (1946). *Probabilidad estadística y verdad*. Madrid: Eppasa Calpe.
- Monkman, M. (1996). Modelling TV audience Data. *Admap*. April. Disponível em <http://www.warc.com/Pages/Taxonomy/Results.aspx?q=&Area=&Page=7&Tab=&DVals=&SourceOR=&DRange=&Filter=All&SubjectRef=457>
- Moore, A., Schneider, J., & Deng, K. (1997). Efficient locally weighted polynomial regression predictions. In D. Fisher (Ed.), *Proceedings of the 14th International Conference on Machine Learning (ICML-97)*. Burlington, Massachusetts: Morgan Kaufmann Publishers.
- Mora, J., Jason, H., & Krider, R. (2011). Television co-viewing in Mexico: An assessment on the people meter data. *Journal of Broadcasting & Electronic Media*, 55(4), 448–469.
- Moritz, G., & Pereira M. (2006). *Processo decisório*, (1ª edição). Florianópolis: UFSC.
- Morley, D. (1996). TV, audiencias y estudios culturales. Buenos Aires: Amorrortu.
- Morchen, F., & Ultsch, A. (2004). Discovering temporal knowledge in multivariate time series. In *Proceedings of the 28th Annual Conference of the Gesellschaft für Klassifikation e.V. University of Dortmund, Germany, 9-11 March 2004* (pp.272-279).
- Morley, D. (1998). Populismo, revisionismo y los nuevos estudios de la audiencia. In J. Curran, D. Morley & V. Wakerdine (Eds). *Estudios culturales y comunicación* (p. 417-438). Barcelona: Paidós.
- Murphy, J., Cunningham, I., & Wilcox, G. (1979). The impact of program environment on recall of humorous television commercials. *Journal of Advertising*, 8(2), 17-21.
- Murray, G. & Jenkins, J. (1992). The concept of “Effective Reach” in advertising, *Journal of Advertising Research*, 32(3), 34-42.
- Murteira, B., Ribeiro, C., Silva, J., & Pimenta, C. (2007). *Introdução à estatística*. (2ª ed.), Lisboa: McGraw-Hill.
- Nagel, E. (1974). *La logica sin metafisica*. Madrid: Editorial Tecnos.

- Napoli, P. (2001). The unpredictable audience: An exploratory analysis of forecasting error for new prime-time network television programs. *Journal of Advertising*, 30(2), 53 – 60.
- Neto, E., De Carvalho, F. & Bezerra, L. (2006). Linear regression methods to predict interval-valued data. In *Proceeding of the Ninth Brazilian Symposium on Neural Networks, Ribeiro Preto, Brazil, 23-27 October 2006* (pp.125-130).
- Niblett, T., & Bratko, I. (1986). Learning decision rules in noisy domain. In *Proceedings of the Expert Systems 86*. Cambridge: Cambridge University Press.
- Nienhuys-Cheng, S., & Wolf, R. (1997). *Foundations of inductive logic programming*. New York: Springer-Verlag.
- Nikolopoulos, K. (2010). Forecasting with quantitative methods the impact of special events in time series. *Journal of Applied Economics*, 42(8), 947-955.
- Nikolopoulos, K., Goodwin, P., Patelis, A., & Assinkopoulo, V. (2006). Forecasting with cue information. A comparison of multiple Regression with alternative forecasting approaches. *European Journal of Operation Research*, 180, 354-368.
- Noelle-Neumann, E. (1984). Alcune ricerche di audience. In G. Bettini, *Nuove dimensioni televisive* (pp. 323-330). Turin: Eri/Edizioni RAI.
- Noirhomme-Fraiture, M. (2004). *User manual for SODAS 2 Software*. FUNDP: Belgique, April, version 1.0.
- Nüesch, S., & Franck, E. (2009). The role of patriotism in explaining the TV audience of national team games – Evidence from four international tournaments. *Journal of Media Economics*, 22(6), 6-19.
- Oliver, J. (1993). Decision graphs – an extension of decision trees. In *Proceedings of the fourth International workshop on Artificial Intelligence and Statistics, Fort Lauderdale, Florida, USA, January 1993* (pp.343-350).
- Ostrow, J. (1982). *Setting frequency levels*. In *effective frequency: The state of the art* (pp. 89-102). Key Issues Workshop. New York: Advertising Research Foundation.
- Padhy, N., & Panigrahi, R. (2012). Multi-relational data mining approaches: A data mining technique. *International Journal of Computer Applications* (0975 – 8887), 57(17), 23-32.

- Parsons, T. (1940). An analytical approach to the theory of social stratification. *The American Journal of Sociology*, 45(6), 841-862.
- Patel, S., & Schlijper, A. (2008). *Models of consumer behaviour. Unilever corporate research*. Disponível em <http://www.maths-in-industry.org/past/ESGI/49/problems/unilever2/unilever2.pdf>
- Patelis, A., Metaxiotis, K., Nikolopoulos, K., & Assimakopoulos, V. (2003). FORTV: Decision support system for forecasting television viewership. *Journal of Computer Information Systems*, 43(4), 100–107.
- Patzer, G. (1991). Multiple dimensions of performance for 30 and 15 second commercials. *Journal of Advertising Research*, 31(4), 18-25.
- Pechman, C. & Stewart, D. (1988). Advertising repetition: A critical review of wearin and wearout. *Current Issues and Research in Advertising*, 11, 285-329.
- Pingree, S., Hawkins, R., Hitchon, J., Giligan, E., Radler, B., Kahlor, L., Gorham, B., Kolbeins, G., Schmidt, T., & Kannaovakun, P. (2001). If college students are appointment television viewers. *Journal of Broadcasting and Electronic Media*, 45(3), 446-463.
- Plessis, E. (1994). Recognition versus recall. *Journal of Advertising Research*, 34(3), 75-91.
- Poincaré, H. (1990). Chance. In M. Hitchins, M. Adler & C. Fadiman (Eds). *Gateway to the great books*. Chicago: Encyclopaedia Britannica.
- Poggio, T. & Girosi, F. (1990). Networks for approximation and learning. In *Proceedings of the Institute of Electrical and Electronics Engineers (IEEE)*. Vol. 78, pp. 1481-1497.
- Popper, K. (1957). *La sociedad abierta y sus enemigos*. Buenos Aires: Paidós.
- Popper (1984). *Post Scriptum a la lógica de la investigación científica*, Vol II, *El Universo Abierto: Un argumento a favour del indeterminismo*. Madrid: Tecnos.
- Popper, K. (2011). *La lógica de la investigación científica*. Madrid: Tecnos.
- Popper, K., & Condry, J. (1995). *Televisão: Um perigo para a democracia*. Lisboa: Gradiva.

- Prakash, B., & Li, L. (2011). Time series clustering: Complex is simpler!. In *Proceedings of the 28th International Conference on Machine Learning, Bellevue, Washington, USA, 28th June - 2nd July 2011* (pp. 185-192).
- Quinlan, J. (1993). *C4.5: Programs for machine learning*. Amsterdam: Morgan Kaufmann Publishers.
- Ramsey, F. (1926), Truth and probability. In R. B. Braithwaite (Ed.), (1931). *The foundations of mathematics and other logical essays* (pp. 156-198). London: Routledge and Kegan Paul.
- Rauch J., & Šimůnek M. (2005). An Alternative Approach to Mining Association Rules. In: T. Y. Lin, S. Ohsuga, C. J. Liao & X. Hu (Eds.): *Foundations and Novel Approaches in Data Mining* (pp. 211-231). Berlin: Springer-Verlag.
- Ray, M., & Webb, P. (1986). Three prescriptions for clutter. *Journal of Advertising Research*, 26(1), 69-77.
- Reddy, S., Aronson, J., & Stam, A. (1998). SPOT: Scheduling programs optimally for television. *Management Science*, 44(1), 83–102.
- Reeth, D. (2011). Television demand for the tour de France: The importance of outcome uncertainty, patriotism and doping. *Hub Research Papers 2011/15. Economics and Management*. September. Disponível em <https://lirias.hubrussel.be/bitstream/123456789/4902/1/11HRP15.pdf>.
- Reichenbach, H. (1935). Über induktion und wahrscheinlichkeit. Bemerkungen zu Karl Poppers logik der forschung. *Erkenntnis*, 5(4), 267-284. Engl. transl. 'Induction and probability: Remarks on Karl Popper's the logic of scientific discovery'. In Reichenbach and Cohen, 1978, Vol. II.
- Reis, E. (1997). *Estatística multivariada aplicada*. Lisboa: Edições Silabo.
- Reyes, R. (2009). *Diccionario Crítico de Ciencias Sociales: (Terminologia científico social)*. Villaviciosa de Odón (Madrid): Plaza y Valdés.
- Robertson, T., Zielinski, J., & Ward, S. (1984). *Consumer behaviour*. New York. Scott, Foresmann and Company.
- Rodriguez, O. (2000). *Classification et modeles lineaires en analyse des donnees symboliques*. (Unpublished doctoral dissertation). Université Paris Dauphine, Paris.

- Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Annals Mathematical Statistics*, 27, 832-837.
- Rosenstein, A., & Grant, A. (1997). Reconceptualizing the role of habit: A new model of television audience activity. *Journal of Broadcasting & Electronic Media*, 41(3), 324-344.
- Rosengren, K., Wenner, L., & Palmgreen, P. (Eds.) (1985). *Media gratifications research: Current perspectives*. Beverly Hills: Sage.
- Rubin, A. (1993). Audience activity and media use. *Communication monographs*, 60 (March), 98-104.
- Rubinson, J. (2009). Empirical effects of advertising effectiveness. *Journal of Advertising Research*, 49(2), 220-226.
- Ruddock, A. (2001). *Understanding audiences. Theory and methods*. Londres: Sage.
- Ruiz, J. (1996). *Metodologia da investigación cualitativa*. Bilbao: Universidade de Deusto.
- Rust, R., & Alpert, M. (1984). An audience flow model of television viewing choice. *Marketing Science*, 3(Spring), 113-124.
- Rust, R., & Eechambadi, N. (1989). Scheduling network television programs: A Heuristic audience flow approach to maximizing audience share. *Journal of Advertising* 18(2), 11-18.
- Rust, R., Kamakura, W., & Alpert, M. (1992). Viewer preference segmentation and viewing choice models for network television. *Journal of Advertising Research*, 21(1), 1-18.
- Ryan, S., & Porth, L. (2007). *A tutorial on the piecewise regression approach applied to bedload transport data*. Gen. Tech. Rep. RMRS-GTR-189. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.
- Salzinger, K. (2005). Clinical, statistical and broken-leg predictions. *Behavior and Philosophy*, 33, 91-99.
- Saperas, E. (1998). *Manual Básico de Teoría de la Comunicación*. Barcelona: CIMS.

- Savage, L. (1954). *The Foundations of Statistics*. (1972 edition). New York: Dover.
- Serra, I. (2005). O indeterminismo e o problema das duas culturas. In *Workshop Indeterminismo*. Disponível em http://cfc.ul.pt/equipa/3_cfc_ul_elegiveis/isabel_serra/isabel%20serra.htm
- Sewell, W. (1992). A theory of structure: Duality, agency, and transformation. *American Journal of Sociology*, 98(1), 1–29.
- Siddarth, S. & Chattopadhyay, A. (1998). To zap or not to zap: A study of the determinants of channel switching during commercials. *Marketing Science*, 17(2), 124-38.
- Silva, E., & Menezes, E. (2005). *Metodologia da pesquisa e elaboração de dissertação* (4ª Ed.). Florianópolis: UFSC.
- Silveira, F. (1993). Determinismo, Previsibilidade e Caos. *Cad. Cat. Ens. Fís.*, 10(2), 137-147.
- Sharp, B., Beal, V., & Collins, M. (2009). Television back to the future. *Journal of Advertising Research*, 49(2), 211-219.
- Sherman, S. (1995). Determinants of repeat viewing to prime time public TV programming. *Journal of Broadcasting & Electronic Media*, 39(4), 472-481.
- Shumann, D., Richard, E., & Scott, D. (1990). Predicting the effectiveness of different strategies of advertising variation: A test of the repetition - Variation hypotheses. *Journal of Consumer Research*, 17(2), 192-202.
- Soares, T. (1996). *SIC versus SIC: A representação social de uma “realidade” televisiva*. Dissertação de Licenciatura em Sociologia, ISCTE, Lisboa, Portugal.
- Sorlin, P. (1997). *Mass media*. Oeiras: Celta Editora.
- Stavenhagen, R. (1973). Estratificação social e estrutura de classes: um ensaio de interpretação. In V. O. Guilherme, M. G. S. Palmeira & A. R. Bertelli (Orgs.), *Estrutura de classes e estratificação social* (pp. 133-70). Rio de Janeiro: Ed. Zahar.
- Stavitsky, A. (2000). Theory into practice. By the numbers: The use of ratings data in academic research. *Journal of Broadcasting & Electronic Media*, 44(3), 535-539.

- Steinmann, M. (1988). Nouveau systeme electronique de mesure de l'audience TV en Suisse-systeme telecontrol de la SSR-Experiences et perspectives. In *Etudes de radio-television* (pp. 37-72). Bruxelles: Radio-Television Beige de la Communaute Française, março, 38.
- Stewart, D. (1989). Measures, methods, and models in advertising research. *Journal of Advertising Research*, 29 (June/July), 54-60.
- Sidney, S. N., & Castellan, J. Jr, (1988). *Non parametric statistics for the behavioural sciences* (2nd Ed.). New York: Mc Graw- Hill International Editions.
- Soong, R. (1988). The statistical reliability of people meter ratings. *Journal of Advertising Research*, 28(1), 50-56.
- Souchon, M. (1990). Les programmeurs et leurs représentation du public. *Réseaux*, 39 (Janvier), 93-108.
- Souchon, M. (1993). Defense et ilustración de la audience. *Research Area*, 5(2).
- Souchon, M. (2003). Pour une utilisation complexe de l'audimétrie. In D. Wolton (Dir.), *L'audience. Presse, radio, télévision, internet* (pp. 157-165). Paris: CNRS Éditions.
- Sueiras, J., Morales, F., & Ibanez, J. (2007). From TV to online advertising: Recent experience from the spanish media. In *Proceedings of Data Mining and Audience Intelligence for Advertising (ADKDD 07)*, San Jose, California, USA, 12 August 2007 (pp. 63-70).
- Taleb, N., & Makridakis, S. (2009). Living in a world of low levels of predictability. *International Journal of Forecasting*, 25(4), 716-733.
- Tang, H. (2007). *Neural networks: Computational models and applications*. Berlin, New York, Springer.
- Takama, Y., & Hattori, S. (2008). Mining association rules from TV watching Log for program recommendation. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 12(1), 26-31.
- Tavares, L., Oliveira, R., Themido, I., & Correia, F. (1996). *Investigação Operacional*, Lisboa: McGraw-Hill.

- Tejada, R. (1995). *Quality of TV publicity impacts*. European Society for Opinion and Marketing Research, Sponsorship and Promotion, Madrid, March.
- Tewksbury, D., Weaver, A., & Maddex, B. (2001). Accidentally informed: Incidental news exposure on the world wide web. *Journalism & Mass Communication Quarterly*, 78(3), 533-554.
- Tiedge, J., & Ksobiech, K. (1986). The Lead-in strategy for prime-time TV: Does it increase the audience?. *Journal of Communication*, 36(2), 51-63.
- Tiedge, J., & Ksobiech, K. (1987). Counterprogramming primetime network television. *Journal of Broadcasting & Electronic Media*, 31(1), 41-55.
- Tille, Y. (1990). Du choix du plan de sondage a la fiabilité des resultats. In J.-J. Drosbeke, & G. Thoverson, *Au royaume des sondages* (pp. 65-70). Bruxelas: Editions de L'Université de Bruxelles.
- The Ad Class (2010). *Super Bowl XLIV 2010 bestads/commercials*. Disponível em <http://www.theadclass.com/events/superbowl-xliv-44-2010-ads>
- Toro, J. (1997). *Análisis del comportamiento de la audiencia en televisión: cuantificación del zapping activo y sus variables condicionantes*. Tesis doctoral, IESE, Universidad de Navarra, Pamplona, España.
- Torres, C. (2004). Televisão do nós e televisão do eu – A encruzilhada da televisão generalista, *Análise Social*, Vol. XXXVIII (169), 1011-1042.
- Torres, A. (2008). A nossa entrada tardia. Participação na Sessão Plenária Mundos Sociais: Saberes e Práticas. Em *VI Congresso Associação Portuguesa de Sociologia*. Lisboa, Portugal, 25-28 de Junho.
- Twyman, T. (1983). *Methoden der zuschauer – und horfunkorschung und ihre validitat*. Agma Schriftenband 8/83. Frankfurt: Media Micro Census.
- Ufman, L. & Rousseeuw, P. (2008). *References, in finding groups in data: An introduction to cluster analysis*. Hoboken, NJ, USA: John Wiley & Sons, Inc. doi: 10.1002/9780470316801.refs.
- Vandebosch, H., Roe, K., & Van den Bulck, J. (2006). Moon and media: Lunar cycles and television viewing. *Media Psychology*, 8(3), 287-299.

- Van der Burg, M., Lauf, E., & Negenborn, R. (2011). *Mediamonitor: The dutch media in 2010*. Hilversum: Commissariaat voor de Media.
- Valente, J., & Chavarrias, I. (2000). Discovering similar patterns in time series. In *Proceeding of Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, Boston, MA, USA, 20-23 August 2000 (pp. 497-505).
- Valêncio, C., Oyama, F., Ichiba, F., & Souza, R. (2011). Multi-relational Algorithm for Mining Association Rules in Large Databases. In *Proceedings of 12th International Conference on Parallel and Distributed Computing, Applications and Technologies*. Gwangju, South Korea, 20 – 22 October 2011 (pp.269-274).
- Van Meurs, L. (1998). Zapp! A study of switching behaviour during commercial breaks. *Journal of Advertising Research*, 38(1), 43-53.
- Venn, J. (1866). *The logic of chance: An essay on the foundations and province of the theory of probability*. First Edition. London and Cambridge: Macmillan and Co.
- Vicente, P., Reis, E., & Ferrão, F. (1996). *Sondagens. A amostragem como factor decisivo da qualidade*. Lisboa: Edições Silabo.
- Vos, B., Irazoqui, M., Nikkel, J., & Hoogendoorn, A. (2009). *How reliable in my audience? Coping with audience segmentation*. ESOMAR, Worldwide Media Measurement, Stockholm, 4-6 May.
- Wakshlag, J. (1985). Repeat viewing in Boston panel. In *Transcript Proceedings: Advertising research foundation electronic media and research technologies-IV*. New York: Advertising Research Foundation.
- Walker, J. (1988). Inheritance effects in the new media environment. *Journal of Broadcasting & Electronic Media*, 32(4), 391-401.
- Wartofsky, W. (1973). *Introducción a la filosofía de la ciencia*. Madrid: Alianza Universidad.
- Weber, R. (2002). Methods to forecast television viewing patterns for target audiences. In Schorr A., Campbell B., & Schenk M., *Communication research in Europe and abroad – Challenges of the first decade*. Berlin: DeGruyter.
- Webster, J. (1985). Program audience duplication: A study of television inheritance effects. *Journal of Broadcasting & Electronic Media*, 29(2), 121-133.

- Webster, J. (2005). Beneath the veneer of fragmentation: Television audience polarization in a multichannel world. *Journal of Communication*, 55(2), 366–382.
- Webster, J. (2006). Audience flow past and present: television inheritance effects reconsidered. *Journal of Broadcasting and Electronic Media*, 50(2), 323-337.
- Webster, J., & Lichty, L. (1991). *Ratings analysis: Theory and practice*. Hillsdale, NJ: Lawrence Erlbaum.
- Webster, J., & Wakshlag, J. (1983). A theory of television program choice. *Communication Research*, 10, 430-446.
- Webster, J., & Wang, T. (1992). Structural determinants of exposure to TV: The case of repeat viewing. *Journal of Broadcasting & Electronic Media*, 26(2), 125-136.
- Wei-Yin Loh (2011). Classification and regression trees. *WIREs Data Mining Knowledge Discovery*, 1(1), 14-23. doi: 10.1002/widm.8.
- Wert, J. (1990). Todo lo que usted quiso saber sobre el teléfono y no-se atrevio a perguntar, *Anuncios*, 341, 42-43.
- Whittaker, E. (1943). Chance, freewill and necessity in the scientific conception of the universe. *Proceedings of the Physical Society*, 55(6), 459-71.
- Wickelmaier, F. (2003). *An introduction to MDS*. Sound Quality Research Unit, Aalborg University, Denmark.
- Wiu, Y., Agrawal, D., & Abaddi, A. (2000). A comparison of DFT and DWT based similarity search in time series databases. In *Proceedings of 9th International Conerence on Information and Knowledge Management (CIKM)*, McLean, VA, USA, 6-11 November 2000 (pp. 488-495).
- Wolf, M. (1991). *La investigación de la comunicación de massas. Critica y perspectivas*. Barcelona: Paidós.
- Wonneberger, A., Schoenbach, K., & Van Meurs, L. (2011). Interest in news and politics or situational determinants? Why people watch the news. *Journal of Broadcasting & Electronic Media*, 55(3), 325-343.

- Wonneberger, A., Schoenbach, K., & Van Meurs, L. (2012). Staying tuned: Tv news audiences in the Netherlands: 1988-2010. *Journal of Broadcasting & Electronic Media*, 56(1), 55-74.
- Wulf, P., Verhagen, E., & Warty, M. (2006). Towards better predictive programming – generating engagement patterns. *ESOMAR, Annual Congress*. London, 17-20 de setembro.
- Wyndham, J. (1994). Strategic Development Director. MRB Grop. Board Director. BMRB Internacional and Brenan, D., Controller, Audience Planning. ITV Center. United Kingdom. A Series of Success? Modelling and Predicting the Success of a TV program Series.
- Xiaoling, L. (2007). *Modelling of television audience viewing behaviour*. Ph.D Thesis, City University of Hong Kong, Hong Kong, China.
- Xiaoling, L., & Hing-Po, L. (2006). Forecasting the number of audience of TV programs: A bayesian approach. Communication at the *26th International Symposium on Forecasting*, Santander, Spain.
- Xiaoling, L., & Hing-Po, L. (2007). Repeat viewing and viewing intensity of TV Programs. Communication at the *6th International Conference on Information and Management Sciences*, Lhasa, China.
- Xin, D., Han, J., Cheng, H., & Yan, X. (2007) Frequent pattern mining: current status and future directions, *Data Mining Knowledge Discovery*, 15(1), 55–86.
- Yang, K. (2012). Multivariate local polynomial regression with autocorrelated errors. *Economics Bulletin, AccessEcon*, 32(4), 3298-3305.
- Yang, M., Hung, W., & Chen, D. (2012). Self-organizing map for symbolic data. *Fuzzy Sets and Systems*, 203, 49–73.
- Yao, D. (2009). Global ads forecast to fall 6.9% in 2009. Disponível em <http://news.smh.com.au/breaking-news-business/global-adsforecast-to-fall-69-in-2009-20090415-a6k5.html>
- Yoo, C., & Kim, H. (2002). An analysis of prediction error for new prime-time television programmes: A competitive study between USA and Korea. *International Journal of Advertising*, 21, 525–546.

- Yuan, E., & Ksiazek, T. (2011). The duality of structure in China's national television Market: A network analysis of audience behaviour. *Journal of Broadcasting & Electronic Media*. 55(2), 180–197.
- Zaki, M. (2001). Spade: An efficient algorithm for mining frequent sequences. *Machine Learning Journal*, 42(1/2), 31–60.
- Zarate, M. (1997). *Los modelos de predicción de las audiencias de los programas*. Tesis Doctoral, Universidad Complutense Madrid, Madrid, España.
- ZenithOptimedia (2009). ZenithOptimedia releases April 2009ad expenditure forecasts. Disponível em <http://www.zawya.com/story.cfm/sidZAWYA20090414132254/%20ZenithOptimedia%20releases%20April%202009%20ad%20expenditure%20forecasts>
- Zubayr, C. (1999). The loyal viewer? Patterns of repeat viewing in Germany. *Journal of Broadcasting & Electronic Media*, 4(3), 346-363.
- Zufryden, F. (1973). Media scheduling: A stochastic dynamic model approach. *Management Science*, 19(12), 1395-1406.

Cristina Rosalinda Mendes Ventura

**Os Modelos de Exposição Necessários à Aquisição de Publicidade no
Sector Televisivo**

APÊNDICES

**Universidade Fernando Pessoa
Porto 2015**

ÍNDICE TEMÁTICO

ÍNDICE.....	iii
ÍNDICE TEMÁTICO.....	iii
ÍNDICE DE FIGURAS.....	ix
ÍNDICE DE QUADROS.....	xi
ÍNDICE DE GRÁFICOS.....	xv
ÍNDICE DE TABELAS.....	xix
 APÊNDICE A. DESCRIÇÃO E CRÍTICA AOS MÉTODOS DE MEDIÇÃO DAS AUDIÊNCIAS TELEVISIVAS.....	 1
A.1. INTRODUÇÃO AOS MÉTODOS DE MEDIÇÃO DAS AUDIÊNCIAS.....	1
A.1.1. Os estudos de memória.....	4
A.1.2. Os estudos coincidentais.....	5
A.1.3. Os diários de recepção.....	6
A.1.4. A audimetria.....	7
A.1.5. O audio-matching.....	17
A.2. A DESCRIÇÃO DOS INSTRUMENTOS E PROCEDIMENTOS DE RECOLHA E ESTRUTURAÇÃO DOS DADOS UTILIZADOS.....	18
A.2.1. A descrição do processo de audimetria.....	18
A.2.2. A descrição do painel de audimetria.....	20
A.2.3. O método de amostragem.....	26
A.2.4. A caracterização das bases de dados.....	28
A.2.4.1. A caracterização da base de dados Telereport.....	28
A.2.4.2. A caracterização da base de dados Publitex.....	32
A.2.5. A validade e adequabilidade das fontes de dados.....	33

APÊNDICE B. A INDÚSTRIA TELEVISIVA E A ACTIVIDADE DE PLANEAMENTO DE <i>MEDIA</i>	39
B.1. CARACTERIZAÇÃO DO SECTOR TELEVISIVO NACIONAL.....	39
B.1.1. As estações televisivas e a programação.....	40
B.1.2. As estações televisivas e a programação.....	42
B.1.3. A evolução dos shares de audiências dos canais generalistas.....	44
B.1.4. Os canais por Tv Cabo em Portugal.....	47
B.1.5. As perspectivas futuras para o sector televisivo	49
B.2. A TELEVISÃO COMO MEIO PUBLICITÁRIO	51
B.2.1. Os motivos para valorizar o investimento publicitário no sector televisivo	51
B.2.2. Os factores e possibilidades de negociação na TV: As condições de contratação	52
B.3. A ACTIVIDADE DE PLANEAMENTO DE <i>MEDIA</i>	55
B.3.1. O planeamento de <i>media</i> e os “estudos de audiências”	55
B.3.2. A evolução dos métodos de medição das audiências.....	56
B.3.3. Os pressupostos em que se baseia a investigação das audiências pré-audimetria	57
B.3.4. O contributo da audimetria para o planeamento de <i>media</i>	57
B.3.5. A elaboração de uma campanha publicitária.....	58
B.3.5.1. Os actores e as suas funções	58
B.3.5.2. As etapas e a responsabilidade de uma campanha.....	58
B.3.5.3. A estratégia de <i>media</i>	59
B.3.5.4. O plano de <i>media</i> : A táctica.....	60
B.3.5.5. Os modelos de <i>media</i> : A exposição, a optimização e a avaliação dos efeitos da publicidade	63
B.3.5.6. A necessidade de desenvolver análises explicativas e de previsão das audiências	64
APÊNDICE C. A INVESTIGAÇÃO REALIZADA NO SECTOR TELEVISIVO E ENQUADRAMENTO DA TEMÁTICA PROPOSTA	67
C.1. AS ORIENTAÇÕES TEÓRICAS, OBJECTOS, DISCIPLINAS E ÁREAS DE ANÁLISE.....	67

C.2. AS FONTES DE DADOS UTILIZADAS PARA A ANÁLISE DA EXPOSIÇÃO TELEVISIVA.....	71
C.2.1. Os instrumentos de registro: A audimetria e os “ratings”	72
C.2.2. Os inquéritos e as entrevistas.....	75
C.3. OS MÉTODOS DE INVESTIGAÇÃO APLICADOS AOS ESTUDOS DE AUDIÊNCIAS.....	76
C.3.1. A finalidade da investigação	76
C.3.2. A metodologia quantitativa e qualitativa	78
C.3.3. As abordagens analíticas	80
C.4. A INVESTIGAÇÃO ACADÉMICA E A EMPRESARIAL.....	84
C.5. ENQUADRAMENTO DAS ANÁLISES A REALIZAR.....	85
APÊNDICE D. PERSPECTIVA DOS MODELOS EXPLICATIVOS DO COMPORTAMENTO DO TELESPECTADOR.....	91
D.1. A ABORDAGEM TEÓRICA DO COMPORTAMENTO DO CONSUMIDOR.....	91
D.2. AS PERSPECTIVAS TEÓRICAS QUE ORIENTAM A EXPLICAÇÃO DA EXPOSIÇÃO TELEVISIVA	92
D.2.1. A perspectiva das gratificações e usos do <i>media</i>	94
D.2.2. A abordagem estrutural.....	95
D.2.3. A comparação entre ambas as perspectivas	96
D.3. A AUDIMETRIA E AS TEORIAS QUE EXPLICAM A EXPOSIÇÃO TELEVISIVA.....	101
D.4. AS ANÁLISES EXPLICATIVAS E AS ANÁLISES DE PREVISÃO DA EXPOSIÇÃO TELEVISIVA.....	105
APÊNDICE E. PERSPECTIVA DAS ANÁLISES DE PREVISÃO NAS CIÊNCIAS SOCIAIS.....	109
E.1. CONSIDERAÇÕES A RESPEITO DAS ANÁLISES DE PREVISÃO NAS CIÊNCIAS SOCIAIS	109

E.1.1. As análises de previsão nas ciências sociais e as suas implicações na área da audimetria	109
E.1.2. Considerações sobre os métodos científicos nas ciências sociais.....	117
E.1.3. A indução, a inferência e as abordagens metodológicas	118
E.1.4. A abrangência do conceito de “previsão”	124
E.1.5. Algumas dificuldades que a previsão coloca nas ciências sociais.....	126
 APÊNDICE F. APRESENTAÇÃO DE RESULTADOS	135
F.1. ANÁLISES EXPLORATÓRIAS AOS INDICADORES DE AUDIMETRIA QUE CONDICIONAM A CONSTRUÇÃO DOS PLANOS DE MEDIA.....	139
F.1.1. Análise de taxas médias de audiências para o total da Televisão	139
F.1.1.1. Introdução	140
F.1.1.2. Objectivos	151
F.1.1.3. Bases de dados.....	153
F.1.1.4. Metodologias	156
F.1.1.5. Resultados	158
F.1.2. Análises exploratórias às taxas médias de audiências e aos valores de <i>shares</i> de audiências	188
F.1.2.1. Introdução	188
F.1.2.2. Objectivos	192
F.1.2.3. Base de dados.....	194
F.1.2.4. Metodologias	197
F.1.2.5. Resultados	198
F.2. ANÁLISES EXPLORATÓRIAS DE PREVISÃO DOS INDICADORES DE AUDIMETRIA QUE CONDICIONAM A CONSTRUÇÃO DOS PLANOS DE MEDIA	219
F.2.1. Análises exploratórias de previsão das taxas médias de audiências para o total da TV	220
F.2.1.1. Introdução e objectivos.....	220
F.2.1.2. Dados e metodologias.....	221
F.2.1.3. Resultados.....	224

F.2.2. Análises exploratórias de previsão de valores de <i>Ratings</i>	235
F.2.2.1. Objectivos.....	235
F.2.2.2. Dados e metodologias.....	236
F.2.2.3. Resultados.....	239
F.3. ANÁLISE DE INDICADORES DE DUPLICAÇÃO DAS AUDIÊNCIAS: A LEALDADE PROGRAMÁTICA.....	253
F.3.1. Introdução	253
F.3.2. Objectivos e hipóteses	258
F.3.3. As bases de dados e metodologias	261
F.3.5. Resultados.....	265
F.4. ANÁLISE DE ÍNDICES DE DUPLICAÇÃO DAS AUDIÊNCIAS ENTRE PROGRAMAS E FAIXAS HORÁRIAS.....	272
F.4.1. Introdução	272
F.4.2. Objectivos	282
F.4.3. Bases de Dados.....	288
F.4.4. Metodologia	293
F.4.5. Resultados.....	295
F.5. A SELECÇÃO DOS SUPORTES A INCLUIR EM PLANOS DE <i>MEDIA</i>	327
F.5.1. Introdução	327
F.5.2. Objectivos	333
F.5.3. Bases de dados	335
F.5.4. Metodologias.....	337
F.5.5. Resultados.....	344
F.6. A EXTRACÇÃO DE REGULARIDADES ESTATÍSTICAS: AS ANÁLISES EXPLICATIVAS E DE PREVISÃO MULTIVARIADAS DE VALORES DE RATINGS PROGRAMÁTICOS.....	367
F.6.1. Introdução	367
F.6.2. Objectivos.....	368

F.6.3. Metodologias e bases de dados	370
F.6.4. Resultados.....	375
F.7. A EXTRACÇÃO DE REGULARIDADES ESTATÍSTICAS: AS ANÁLISES EXPLICATIVAS E DE PREVISÃO DAS AUDIÊNCIAS COMERCIAIS.....	414
F.7.1. Introdução	414
F.7.2. Objectivos	417
F.7.3. Bases de dados	419
F.7.4. Metodologias.....	422
F.7.5. Resultados.....	428
APÊNDICE G. ÁREAS METODOLÓGICAS.....	485
G.1 INTRODUÇÃO.....	485
G.1.1. Comparação dos algoritmos CHAID e CART de Árvores de Regressão.....	487
G.1.2. Comparação de árvores de regressão e a regressão linear múltipla.....	489
G.1.3. Comparação das redes neuronais RBF e MLP.....	492

ÍNDICE DE FIGURAS

Fig. A.2.4.1.1 – Tabela de programação do software Telereport. Fonte: Site Institucional da Marktest.....	28
Fig. F.1.1.3.1 – Estruturação das bases de dados associadas à variável TVR%. Fonte: Elaborado pela autora.....	153
Fig. F.1.2.3.1 – Estruturação de base de dados de <i>Ratings</i>	186
Fig. F.2.1.3.III.1 – Segmentos horários que resultam da aplicação de medidas de dissimilaridade simbólicas (SO1).....	230
Fig. F.2.1.3.III.2 – Segmentos horários para os dados de fim-de-semana (medida SO4).....	231
Fig. F.2.1.3.III.3 – Segmentos horários para o período semanal (SO4).....	232
Fig. F.2.1.3.III.4 – Resultado da aplicação do algoritmo de <i>Clustering</i> Divisivo a <i>Tvr_Período%</i> (solução a 5 <i>clusters</i>).....	233
Fig. F.2.1.3.III.5 – Representação de <i>clusters</i> horários para o período semanal e de fim-de-semana.....	234
Fig. F.2.2.3.II.1 – Representação de segmentos horários para o canal RTP1.....	243
Fig. F.2.2.3.II.2 – Representação de <i>clusters</i> horários para o canal TVI.....	244
Fig. F.2.2.3.II.3 – Representação dos <i>clusters</i> horários para o canal SIC.....	245
Fig. F.2.2.3.II.4 – <i>Clustering</i> divisivo dos períodos horários (solução para 4 <i>clusters</i>).....	248
Fig. F.4.5.III.1 – Resultados da análise MDS para as dimensões 1,2 e 3.....	301
Fig. F.6.4.1 – Representação da árvore de regressão CART.....	396
Fig. F.6.4.2 – Pesos da Rede Neuronal MLP.....	411
Fig. F.6.4.3 – Pesos da Rede Neuronal MLP.....	411
Fig. F.7.3.1 – Caracterização da construção das bases de dados de audiências comerciais.....	419

Fig. F.7.4.1 – Metodologias a utilizar nas análises comerciais. Fonte: Elaborado pela autora.....	422
Fig. F.7.4.2 – Tipologias de variáveis simbólicas.....	424

ÍNDICE DE QUADROS

Quadro A.2.2.1 – Caracterização das variáveis de segmentação do painel. Fonte: Site institucional da Marktest.....	21
Quadro A.2.2.2 – As classes sociais Marktest. Fonte: Site institucional da Marktest.....	22
Quadro A.2.2.3 – Caracterização das classes sociais (Ocupação / Nível de Estudos). Fonte: Agência de meios Tempo-Omd.....	23
Quadro A.2.5.1 – Vantagens e desvantagens dos métodos de amostragem probabilísticos e não probabilísticos. Fonte: ALEA - Acção local estatística aplicada.....	37
Quadro B.1.1.1 – A evolução da programação. Fonte: Adaptado de CRECS (2006) e ERC (2010).....	41
Quadro B.1.3.1 – Evolução de <i>Shares</i> (1992-2010). Fonte: Marktest (2009).....	44
Quadro B.1.4.1 – Evolução do número de alojamentos e do número de assinantes de TV por Cabo (1995-2010). Valores em milhares. Fonte: Anacom (2010).....	47
Quadro B.3.5.4.4.1 – Etapas da estratégias de meios. Fonte: Elaborado pela autora.....	62
Quadro C.5.1 – Um mapa da investigação na área televisiva. Fonte: Casetti & Chio (1997).....	89
Quadro D.2.3.1 – Teoria de gratificações e usos do <i>media</i> e teoria estrutural da exposição televisiva. Fonte: Elaborado pela autora.....	99
Quadro D.2.3.2 – Teoria de gratificações e usos e teoria estrutural da exposição televisiva. Fonte: Elaborado pela autora.....	100
Quadro E.1.3.1 – Metodologias deterministas e probabilísticos. Fonte: Elaborado pela autora.....	120
Quadro F.1 – Objectivos das análises de duplicação das audiências. Fonte: Elaborado pela autora.....	136
Quadro F.1.1.5.III.1 – Identificação de segmentos mais coincidentes na sua exposição televisiva.	176

Quadro F.1.2.3.1 – Cenários de exposição televisiva e cálculo dos valores de <i>ratings</i>	195
Quadro F.2.1.3.III.1 – Base de dados simbólica para a análise de <i>Tvr_Periodo%</i>	229
Quadro F.2.1.3.III.2 – Aplicação do algoritmo DIV para a identificação de segmentos horários.....	232
Quadro F.2.2.2.1 – As metodologias de <i>clustering</i> de séries temporais de valores de <i>ratings</i> adoptadas.....	237
Quadro F.2.2.2.2 – A estrutura metodológica das análises propostas para a análise dos <i>ratings</i>	238
Quadro F.3.2.1 – Hipóteses a testar quanto à repercussão de factores estruturais na duplicação da audiência. Fonte: Elaborado pela autora.....	259
Quadro F.3.3.1 – Índices de duplicação das audiências de lealdade programática. Fonte: Elaborado pela autora.....	261
Quadro F.3.3.2 – Bases de dados para a análise de índices de lealdade programática. Fonte: Elaborada pela autora.....	262
Quadro F.4.3.1 – Perspectivas de quantificação da duplicação das audiências que permitem avaliar os efeitos de herança entre os suportes publicitários.....	288
Quadro F.4.4.2 – Índices de duplicação de audiências (<i>software</i> Telereport).....	289
Quadro F.4.3.3 – Base de dados de índices de duplicação das audiências.....	289
Quadro F.4.3.4 – Base de dados de índices de duplicação para as faixas horárias.....	290
Quadro F.4.3.5 – Índices de fluxos de audiências entre os suportes horários ou programáticos.....	291
Quadro F.4.3.6 – Base de dados de fluxos de audiências.....	292
Quadro F.4.5.II.1 – Média de duplicação de audiências em função da posição dos programas para o canal SIC.....	297
Quadro F.5.3.1 – Caracterização das variáveis que traduzem a exposição sociodemográfica e que viabilizam a selecção dos melhores suportes a incluir nos planos de <i>Media</i>	336
Quadro F.5.4.1 – Metodologias utilizadas nas análises de associação de suportes a segmentos sociodemográficos.....	338
Quadro F.5.4.2 – Metodologias que contribuem para a segmentação das audiências televisivas.....	343
Quadro F.5.5.I.1 – Identificação de alguns números que constam nos Mapas Perceptuais.....	349

Quadro F.5.5.I.2 – Associação de programas a faixas etárias de acordo com a aplicação de metodologias MDS.....	352
Quadro F.5.5.I.3 – Identificação de programas utilizados na análise de <i>cluster</i>	353
Quadro F.5.5.II.1 – Programas do canal RTP1 com menores índices de adesão.....	360
Quadro F.5.5.II.2 – Programas do canal RTP1 com maiores índices de adesão.....	360
Quadro F.5.5.II.3 – Programas do canal TVI com maiores e menores índices de adesão.....	361
Quadro F.5.5.II.4 – Programas do canal SIC com maiores e menores índices de adesão.....	362
Quadro F.6.3.1 – Descrição de algumas das variáveis que constam das análises multivariadas a realizar.....	371
Quadro F.7.1.1 – Áreas temáticas no âmbito da investigação das audiências comerciais.....	415
Quadro F.7.3.1 – Descrição da base de dados.....	421
Quadro G.1 – Áreas Metodológicas.....	485

ÍNDICE DE GRÁFICOS

Gráfico F.1.1.5.I.1 – Evolução da variável <i>TVR_Prog%</i> associada ao período semanal e de fim-de-semana.....	159
Gráfico F.1.1.5.III.1 – Evoluções médias de <i>Tvr_Período%</i> para os segmentos que apresentam maior ou menor exposição.....	177
Gráfico F.1.1.5.IV.1 – Evolução de <i>Tvr_Período%</i> para os espaços comerciais.....	177
Gráfico F.1.2.1.1 – Evolução dos indicadores <i>Tvr%</i> , <i>Rat%</i> e <i>Shr%</i> considerando o universo televisivo.....	189
Gráfico F.1.2.5.IV.1 – Aplicação da medida U1 de dissimilaridade entre SO booleanos aos dados de fim-de-semana do canal SIC e que permite uma análise da variabilidade dos <i>ratings</i> sociodemográficos.....	214
Gráfico F.1.2.5.IV.2 – Aplicação da medida U3 aos dados semanais do canal SIC e que permite uma análise da variabilidade dos <i>ratings</i> sociodemográficos.....	214
Gráfico F.2.1.3.I.1 – Previsão com uma regressão ultra-local linear.....	226
Gráfico F.2.1.3.I.2 – Representação dos valores de resíduos da regressão Local Linear.....	226
Gráfico F.2.1.3.I.3 – Previsão com uma regressão kernel com 6 KNN.....	226
Gráfico F.2.1.3.I.4 – Representação dos valores de resíduos da Regressão Kernel com 6KNN.....	226
Gráfico F.2.2.3.I.1 – O <i>clustering</i> de séries temporais de valores de <i>ratings</i> para o canal SIC.....	239
Gráfico F.2.2.3.I.2 – O <i>Clustering</i> de séries temporais de valores de <i>ratings</i> para o canal TVI.....	240
Gráfico F.2.2.3.I.3 – O <i>Clustering</i> de séries temporais para o canal RTP1.....	240
Gráfico F.4.5.III.2 – A duplicação de audiências face o período 19/1Q.....	303
Gráfico F.4.5.III.3 – A duplicação de audiências face o período 20/1Q.....	303
Gráfico F.4.5.III.4 – A duplicação de audiências face o período 22/1Q.....	303

Gráfico F.4.5.III.5 – A duplicação de audiências face o período 21/1Q.....	303
Gráfico F.4.5.III.6 – A duplicação de audiências face ao período 23/1Q.....	304
Gráfico F.4.5.III.7 – Índices de duplicação entre períodos consecutivos.....	304
Gráfico F.4.5.III.8 – Limites de variabilidade para os valores de fluxos face ao período 19/1Q.....	307
Gráfico F.4.5.III.9 – Limites de variabilidade para os valores de fluxos face ao período 23/1Q.....	307
Gráfico F.4.5.III.10 – Limites de variabilidade para os valores de fluxos face ao período 0/1Q.....	307
Gráfico F.4.5.III.11 – Limites de variabilidade para os valores de fluxos face ao período 21/1Q.....	307
Gráfico F.4.5.III.12 – Evolução de Tv_On e Tv_Off.....	306
Gráfico F.4.5.III.13 – Índices de duplicação médios face ao período 13/1Q.....	308
Gráfico F.4.5.III.14 – Variabilidade de índices de duplicação face ao período 13/1Q.....	308
Gráfico F.5.5.I.1 – Correlações das variáveis com os eixos para a ACP Simbólica realizada com os programas da TVI (idêntico à ACP clássica, algoritmo dos vértices).....	345
Gráfico F.5.5.I.2 – Mapa perceptual para as classes sociais e programas do canal TVI (Dim. 1 e 2).....	348
Gráfico F.5.5.I.3 – Mapa perceptual para as classes sociais e programas do Canal TVI (Dim. 2 e 3).....	348
Gráfico F.5.5.I.4 – Variáveis estatisticamente significativas na regressão com as dimensões MDS.....	350
Gráfico F.5.5.I.5 – Intervalo de correlações das variáveis segundo o algoritmo dos vértices.....	356
Gráfico F.5.5.I.6 – Classes sociais no espaço MDS.....	356
Gráfico F.5.5.I.7 – Variáveis significativas na regressão (faixas etárias).....	357
Gráfico F.5.5.III.1 – Correlação entre índices de afinidade para as regiões geográficas.....	364
Gráfico F.5.5.III.2 – Correlação entre índices de afinidade para as classes sociais.....	364

Gráfico F.5.5.III.2 – Correlação entre índices de afinidade para as classes sociais.....	364
Gráfico F.5.5.III.3 – Correlação entre índices de afinidade para as faixas etárias.....	365
Gráfico F.5.5.III.4 – Correlação entre índices de afinidade para o sexo e situação no lar.....	365
Gráfico F.5.5.III.5 – Correlação entre índices de afinidade para a totalidade dos segmentos.....	365
Gráfico F.7.5.1 – Evolução de <i>ratings</i> comerciais para o canal SIC.....	435
Gráfico F.7.5.2 – Valores previstos e observados.....	445
Gráfico F.7.5.3 – Valores observados e previstos.....	447

ÍNDICE DE TABELAS

Tabela F.1.1.5.II.1 – Variabilidade de valores da variável <i>TVR_Periodo%</i>	171
Tabela F.1.1.5.II.2 – Variabilidade de valores da variável <i>TVR_Periodo%</i>	171
Tabela F.2.1.3.III.1 – Descrição da constituição dos segmentos horários.....	234
Tabela F.2.2.3.II.1 – Representação de faixas horários por ordem crescente de valores de <i>Rat_Período%</i>	245
Tabela F.1.2.2.3.II.2 – Valores de <i>ratings</i> para cada segmento horário identificado com as Redes SOM simbólicas.....	247
Tabela F.4.5.IV.1 – Correlação entre <i>ratings</i> de pares de programas adjacentes sob elevada e reduzida duplicação primária de audiências.....	319
Tabela F.5.5.I.1 – Interpretação das Dimensões dos Mapas Perceptuais.....	348
Tabela F.5.5.I.2 – Interpretação das Dimensões para o MDS que associa programas a faixas etárias.....	350
Tabela F.5.5.I.3 – Matriz de Classificações.....	351
Tabela F.5.5.I.4 – Coeficientes Estruturais das Variáveis com as Funções Discriminantes..	354
Tabela F.5.5.I.5 – Coeficientes das Variáveis Sexo e Situação no Lar.....	357
Tabela F.5.5.III.1 – Descrição de factores.....	363
Tabela F.6.4.1 – Capacidade explicativa das Árvores de Regressão.....	409
Tabela F.6.4.2 – Pesos da Rede Neuronal MLP.....	410
Tabela F.6.4.3 – Grau de Ajustamento da Rede Neuronal MLP.....	410
Tabela F.6.4.4 – Pesos da Rede Neuronal MLP.....	411
Tabela F.6.4.5 – Grau de Ajustamento da Rede Neuronal MLP.....	411
Tabela F.6.4.6 – Síntese de capacidade explicativa e de previsão das Redes Neurais.....	413

Tabela F.7.5.1 – Coeficientes de determinação dos Modelos de Regressão Linear Múltipla.....	444
Tabela F.7.5.2 – Coeficientes de Correlação Linear de pearson para o canal SIC.....	434
Tabela F.7.5.3 – Coeficientes de Correlação Linear de pearson para o canal SIC.....	446
Tabela F.7.5.4 – Coeficientes de Correlação para o canal SIC e atendendo à variável <i>Spot_Ratio%</i>	450
Tabela F.7.5.5 – Coeficientes de Correlação para o canal TVI.....	452
Tabela F.7.5.6 – Coeficientes de RLM.....	462
Tabela F.7.5.7 – Variáveis com Betas de RLM simbólica Significativos.....	464
Tabela F.7.5.8 – Resultados das Análises de Regressão Linear para o canal RTP1.....	468
Tabela F.7.5.9 – Resultados das Análises de Regressão Linear para canal RTP1 e <i>Spot_Ratio%</i>	470
Tabela F.7.5.10 – Resultados das análises de Regressão não Linear para o canal SIC.....	472
Tabela F.7.5.11 – Comparação de coeficientes Beta de Regressão e Medidas M do algoritmo CART - Canais privados.....	479
Tabela F.7.5.12 – Comparação de coeficientes Beta de Regressão e Medidas M do algoritmo CART (Canais Privados).....	480
Tabela F.7.5.13 – Comparação de coeficientes Beta de Regressão e Medidas M do algoritmo CART (Canais Estatais).....	480

APÊNDICE A

DESCRIÇÃO E CRÍTICA AOS MÉTODOS DE MEDIÇÃO DAS AUDIÊNCIAS TELEVISIVAS

A audimetria ou o *audio-matching* consistem em tecnologias que possibilitam uma medição rápida, permanente, contínua e pormenorizada das audiências televisivas, convertendo-se, consequentemente nos principais instrumentos de quantificação e previsão das audiências televisivas que existem na actualidade. Ambas as tecnologias proporcionam um conjunto de dados final com as mesmas características, sendo a audimetria no passado a tecnologia adoptada por um maior horizonte temporal no sector televisivo. Neste apêndice demonstra-se as vantagens da audimetria, as quais a convertem numa das melhores alternativas para a caracterização do comportamento dos telespectadores e para a posterior construção dos modelos de *Media*, que condicionam o investimento publicitário. Também, se descrevem e comparam os restantes métodos de medição das audiências televisivas, referindo as suas vantagens e limitações.

A.1. INTRODUÇÃO AOS MÉTODOS DE MEDIÇÃO DAS AUDIÊNCIAS

De acordo com *Huertas (2002)*, a medição das audiências televisivas foi tradicionalmente definida como um processo, que apresenta como principal objectivo, o conhecimento do número de pessoas que assistem aos produtos televisivos, ou seja, prevalece o carácter quantitativo de análise da exposição televisiva. Por conseguinte, segundo a autora, sendo esta a sua principal finalidade, contribuiu para uma utilização indiferenciada da terminologia estudo quantitativo ou medição de audiências quando se pretende denominar este tipo de investigações. Contudo, apesar de incidir na quantificação das audiências, também tornam viável a análise de alguns aspectos qualitativos relacionados com a exposição aos *Media*, mas que, no passado, não foram valorizados (*Huertas, 1997*).

Pontualmente verifica-se este duplo interesse, como, por exemplo, em *Mancisidor (1984)* - onde é referido que estas abordagens também permitem a observação dos hábitos e do comportamento dos telespectadores - e em *Leon (1990)*, onde se salienta que nesta área não se valoriza apenas o número, mas, também, as características dos telespectadores.

Sem dúvida que a medição das audiências televisivas corresponde a um estudo quantitativo, uma vez que, se focaliza em dados mensuráveis, mas permite, simultaneamente, extrair informação quantitativa e qualitativa. O seu principal objectivo consiste em determinar qual é a dimensão do número de telespectadores expostos a determinados suportes publicitários, mas que são classificados a partir de determinadas variáveis, as quais especificam a sua natureza e composição. A análise dos dados permite que sejam retiradas algumas conclusões relativamente aos comportamentos de consumo televisivo, o que em geral corresponde a questões de cariz qualitativo.

Na literatura da especialidade encontramos inúmeras referências bibliográficas que avaliam a adequabilidade dos instrumentos de medição das audiências televisivas, assim como dos diferentes métodos de investigação existentes neste âmbito (*Ghosh, 2008; Gill, 2000; Vos, Irazoqui, Nikkel & Hoogendoorn, 2009*). Na actualidade, os instrumentos que mais se destacam correspondem aos estudos de memória, aos diários de recepção, aos estudos coincidentais, à audimetria e mais recentemente ao “*audio-matching*” e variam consoante o país onde a tecnologia é adoptada.

Nos estudos de memória são realizados inquéritos que visam conhecer as selecções programáticas executadas por um painel. Apresentam grandes limitações, uma vez que são muito onerosos, inviabilizando a possibilidade de serem efectuados com alguma regularidade e em curtos intervalos de tempo. Nos estudos coincidentais obtem-se a informação que se pretende, interrogando o telespectador no decorrer do período em que assiste à televisão, consequentemente as suas limitações técnicas são evidentes. Contudo, os telespectadores apresentam, cada vez mais, uma maior relutância quanto a serem entrevistados.

Inevitavelmente, novas tecnologias foram desenvolvidas para caracterizar o consumo televisivo, nomeadamente a audimetria. O desenvolvimento da audimetria, apesar de onerosa, apresentou grandes vantagens e induziu uma profunda transformação nesta área. A audimetria permitiu uma medição mais rápida, exacta, fiável e simples da audiência, deixando de ser necessário a memorização dos conteúdos programáticos transmitidos. Viabilizou o desenvolvimento de modelos matemáticos que caracterizam diferentes vertentes do consumo televisivo. Os restantes métodos de medição das audiências foram sucessivamente colocados em segundo plano.

No passado, a análise das bases de dados de audimetria nem sempre conduziu a resultados satisfatórios devido à sua dimensão, crescente grau de complexidade dos comportamentos de consumo televisivo e às limitações das ferramentas estatísticas existentes. Os sucessivos avanços observados na área do *Data Mining*, entre outras áreas de análise estatística de grandes bases de dados, podem tornar realizável a extracção de uma maior informação e a construção de modelos que são relevantes para os profissionais do sector.

O contributo dos índices de audiência para o meio televisivo e de radiodifusão é reconhecido mundialmente. Todos os instrumentos que pretendem medir a exposição televisiva apresentam, no entanto, determinadas limitações. Porém, a tecnologia de audimetria ou um sistema de informação equivalente são considerados os melhores processos que existem na actualidade, atendendo às necessidades de informação do sector.

As técnicas de recolha de dados têm sofrido evoluções ao longo do tempo e a investigação tem-se centrado na procura e aperfeiçoamento de determinados aspectos, nomeadamente na capacidade de medir de forma contínua e permanente sobre amostras fixas ou painéis; na capacidade de medir indivíduos, e não apenas lares; na capacidade de reduzir as tarefas que exigem um maior envolvimento por parte do painel, uma vez que quanto maior for a necessidade de intervenção dos telespectadores, maiores também podem ser os problemas que podem surgir neste processo de sondagem.

Uma das desvantagens frequentemente referida a respeito da audimetria corresponde à impossibilidade de conhecer as opiniões dos telespectadores quanto ao que é visionado. É importante salientar que nos diários de audiências esta vertente também não é incluída, apenas nas entrevistas que não são realizadas de forma regular e não constituem um instrumento de previsão. Por outro lado, em alguns países os audímetros permitem que sejam colocadas questões ao painel, através do desenvolvimento da audimetria interactiva.

A.1.1. Os estudos de memória

Os estudos de memória são sondagens periódicas, mas não permanentes, onde são realizados inquéritos que recolhem informação a respeito da exposição televisiva de um painel, no decorrer da última semana e dia a dia. Consequentemente, apresentam inúmeras limitações e é um processo extremamente oneroso quando é realizado com regularidade e em curtos espaços de tempo.

Neste processo pede-se aos telespectadores que reproduzam a sua exposição passada, especificando quais são os programas seleccionados e em que horários. Este método requer um grande esforço de memória que pode deturpar os dados, especialmente se pretendermos recolher informação para reduzidos intervalos de tempo, como, por exemplo, ao nível do quarto de hora. A partir destes dados estendem-se os resultados ao total da população que se pretende medir (*Hyman, 1984*).

A informação recolhida poderá não reflectir o real comportamento do entrevistado, já que o consumo de que se recorda poderá ser excepcional e pouco representativo. Deve-se prestar particular atenção à elaboração e à supervisão de todos os aspectos que possam influenciar toda a investigação realizada por inquérito, quando esta é colocada em prática. Quem integra a amostra poderá contribuir com respostas erróneas de forma intencionada ou involuntária. Para além, da necessidade de serem incluídas as perguntas mais adequadas e que possam garantir o seu bom funcionamento, também é necessário ter em atenção a sua extensão, estrutura, redacção e desenho.

Quando se pretende medir as audiências, a sua extensão aproxima-se dos 30 minutos e dos 40 minutos se a entrevista for presencial; se esta for realizada via telefone dura entre os 18 e os 20 minutos. O cansaço dos entrevistados a nível presencial surge aos 15 minutos e aos 8 minutos se a entrevista for realizada por telefone (*Blankenship, 1979; Wert, 1990*).

A.1.2. Os estudos coincidentais

Por conseguinte, não permite recolher informação considerada fiável. Apelar à memória do painel coloca inúmeras barreiras à realização de uma sondagem, nomeadamente o esquecimento e a confusão dos telespectadores (*Noelle-Neumann, 1984*). Isto é, frequentemente, os telespectadores não se recordam do que viram ou apresentam um efeito “confusão”, o qual se pode agravar com uma crescente oferta televisiva e com a sucessiva mudança dos protagonistas de programa para programa e entre os diversos canais. Por outro lado, a entrevista incide num horizonte temporal muito restrito e ao basear-se na memória dos telespectadores apenas poderá abordar questões que se refiram a um período mais recente e a faixas horárias que apresentem uma maior duração.

Nos estudos coincidentais, os telespectadores são interrogados no exacto momento em que assistem à televisão. Através de um sistema de entrevistas telefónicas, distribuídas ao longo do dia, verifica-se se cada pessoa que integra a amostra tem o televisor ligado, qual é o canal a que está assistir e com quem. Geralmente registam-se os dados que se referem apenas à meia hora que antecede o telefonema. Porém, neste processo, dificilmente se consegue recolher informação que seja considerada fiável quando se pretende a caracterização da exposição televisiva ao segundo ou para breves unidades de tempo e que são fundamentais para o planeamento de *Media*. Apesar de permitir a recolha de dados recentes, assistimos a alguma dificuldade na recolha de informação a respeito da exposição individual para curtos espaços de tempo e no decorrer de todo o dia, semana, mês e ano.

Para se medir a audiência no decorrer de um maior horizonte temporal, dia a dia, seria necessário um elevado número de entrevistadores o que inviabilizaria este processo do ponto de vista económico.

Todavia, também não seria possível verificar se as respostas correspondem à realidade e este estaria dependente da capacidade de memorização do painel. Inicialmente, as entrevistas eram realizadas a nível presencial, quando existiam poucos canais e na sua maioria apenas da televisão pública. Neste contexto, as medições realizadas quatro vezes ao ano, eram consideradas como sendo suficientes para conhecer e quantificar o consumo televisivo, uma vez que se pressupõe a regularidade dos comportamentos.

Com a ampliação da rede telefónica apareceram as entrevistas telefónicas. Contudo, este método foi perdendo vigência como consequência do aparecimento de mecanismos que possibilitam o desenvolvimento de sondagens permanentes sobre amostras fixas, sendo essa a tendência actual. As entrevistas foram colocadas em segundo plano e são utilizadas, apenas, para detectar a fiabilidade e a validade dos métodos alternativos mais actuais e avançados, assim como, nas áreas geográficas onde os sistemas automáticos não são viáveis, quer devido a problemas geográficos, quer devido à sua reduzida extensão, assim como à sua inviabilidade económica. Este processo diferencia-se dos estudos de memória por incidir apenas na meia hora que antecede a entrevista.

A.1.3. Os diários de recepção

Os diários de recepção correspondem a um resumo do consumo televisivo que as famílias que integram a amostra se comprometem a realizar de forma regular. Diariamente, durante um período de 14 dias, estas devem registar a sua exposição televisiva, sendo a unidade de referência o quarto de hora que será atribuído a determinado canal, no caso do telespectador ter assistido em maior grau a esse canal face aos restantes. Os telespectadores devem estar devidamente informados sobre a forma correcta de preencher as grelhas, sendo que o responsável pela investigação tem que acompanhar o responsável de cada agregado familiar e ter em atenção que as anotações têm que ser realizadas da forma correcta. Apenas permite a recolha de dados para quinze dias e o seu tempo de produção demora entre duas a três semanas, ou seja, um período demasiado alargado se o compararmos com as entrevistas telefónicas.

Por um lado, permitem evitar recorrer à memória, mas apresentam as limitações dos inquéritos. Por vezes, os telespectadores esquecem-se de preencher as grelhas no momento em que assistem à TV e a informação só é registada posteriormente e em função do que se recordam. Em regiões onde os telespectadores apresentam uma menor instrução, ou em cidades onde vivem famílias pertencentes a diferenciadas culturas, poderá ocorrer alguns problemas relacionados com a linguagem em que está redigido. Na Austrália (*Mariet, 1990*), foi demonstrado que esta técnica subestima o consumo televisivo entre 20% a 33%. Actualmente é apenas utilizada como um processo secundário de recolha de dados ou como forma de validação da informação.

A.1.4. A audimetria

A totalidade dos instrumentos, anteriormente, descrita não constitui instrumentos de previsão, uma vez que viabiliza, no máximo, o cálculo de uma frequência de exposição que não pode ser quantificada através de um registo que incide num curto espaço de tempo e que é calculada através da informação disponibilizada pelo próprio objecto de medida.

Quando surgiram os audímetros verificou-se uma profunda transformação nos métodos quantitativos de medição das audiências. Intensificou-se o investimento publicitário por ser o *Media* estudado com maior rigor, permitindo aos gestores deter a informação mais fiável. A audimetria viabilizou uma medição rápida, mais exacta e simples da audiência, tornando desnecessária a memorização dos conteúdos programáticos transmitidos. O seu desenvolvimento foi classificado como a primeira revolução observada na área da medição das audiências televisivas (*Agostini, 1988; Durand, 1982; Lu & Kiewit, 1987; Steinmann, 1988*), permitindo uma recolha automática, contínua e permanente de um conjunto de dados que pretendem quantificar a exposição televisiva. A audimetria é utilizada em mais de oitenta países e a sua metodologia assenta em sólidos fundamentos estatísticos aplicados a inúmeros campos do conhecimento actual; a tecnologia que lhe está subjacente actualiza constantemente as funcionalidades e as capacidades dos audímetros.

Nesta tecnologia são considerados uma série de controlos internos e externos para assegurar a qualidade do processo (*Jauset, 2008*). A metodologia foi adoptada, consensualmente, entre as diferentes partes implicadas e impõem-se rigorosos controlos de qualidade. Todos os processos são susceptíveis de serem aperfeiçoados, mas até ao momento não se conhece nenhum sistema alternativo, para além do *audio-marching*, que seja mais eficaz para a avaliação e quantificação das audiências (*Jauset, 2008*).

A indústria televisiva depende dos dados de audimetria, (ou de um conjunto de dados equivalente), desde as estações televisivas, até às agências de planificação de meios e indirectamente todos os que recorrem a este *Media* para divulgar os seus produtos e serviços. Segundo a bibliografia da especialidade, as universidades apenas têm a beneficiar com a sua utilização na investigação académica. Em especial nas universidades Norte Americanas, entre outros países essencialmente anglo-saxónicos, têm surgido nas últimas décadas inúmeros trabalhos que recorrem a esta fonte de dados.

Face aos restantes instrumentos, os dados de audimetria viabilizam a caracterização da exposição televisiva, segundo diversas vertentes, contribuindo fortemente para a análise das entidades e da estrutura da audiência (*Cassetti & Chio, 1997*), isto é, para a análise dos comportamentos do painel com o objectivo de os descrever e captar a sua dimensão quantitativa. Por conseguinte, esta área recorre a instrumentos caracterizados pela sua natureza quantitativa e estatística, como por exemplo, aos inquéritos e à audimetria, mas este último método, pelos motivos já salientados, apresenta grandes mais-valias face aos restantes.

Nos estudos coincidentais interroga-se o telespectador a respeito dos programa que estão a assistir em determinado momento, mas esta informação também é possível obter pelo processo de audimetria, e com um maior rigor e fiabilidade. Dificilmente alguém, através de uma entrevista, facultará a seguinte informação: a hora exacta em que ligou o televisor, quantos minutos assistiu a determinado programa ou a um espaço comercial, quantos minutos após o início de um intervalo optou por um canal concorrente e em que instante regressou ao canal inicial, e assim sucessivamente.

Sendo esta informação necessária para os 365 dias do ano, no decorrer de todo o dia e, ano após ano, seria totalmente inviável de ser calculada recorrendo aos restantes instrumentos de medição das audiências. Através do audímetro, todas as opções e selecções do painel ficam registadas de forma automática, e essa informação chega de forma rápida e contínua a todos os que dependem dela.

Em alguns países, os sistemas de informação têm sido aperfeiçoados, ao ponto de recorrerem a sistemas de infravermelhos para detectarem a presença dos telespectadores em frente ao televisor. Também existe a possibilidade de se considerar números específicos no telecomando, com a função de registar os padrões de consumo televisivo de quem não pertence ao agregado familiar. Tem sido proposto, e já existe, a utilização de audímetros móveis os quais pretendem registar o consumo televisivo do painel fora do lar.

Os estudos de memória interrogam os telespectadores relativamente às suas selecções passadas. Se no próprio momento em que assistem a determinado programa é difícil recolher informação fidedigna quanto à exposição televisiva, muito mais difícil será relativamente às selecções históricas. Se a nível internacional a audimetria foi adoptada, apesar dos elevados custos que lhe estão associados, certamente que é devido às suas mais-valias, face aos métodos anteriormente existentes. A audimetria permite a recolha de informação que é muito mais fiável, precisa, objectiva, contínua, abundante e rápida do que a que resultaria de interrogar o telespectador a respeito dos seus comportamentos, independentemente do método utilizado.

A informação disponibilizada é mais precisa, uma vez que se obtêm dados que registam a exposição ao segundo, não existindo outro modo de avaliar o consumo televisivo em unidades de tempo tão reduzidas. Utiliza-se um painel ou uma amostra fixa, permitindo a observação e o estudo dos comportamentos individuais durante longos horizontes temporais (fidelidade). Não existe a possibilidade de erro, típica dos diários de consumo, cuja informação se baseia na memória (objectividade). Em pouco tempo, aproximadamente 6 horas, processam-se todos os dados e rapidamente ficam disponíveis os resultados para a sua análise e posterior toma de decisões (rapidez).

Corresponde a uma fonte de dados contínua, durante 24 horas e 365 dias do ano, que permite a recolha de um vasto conjunto de dados, os quais caracterizam a exposição televisiva sobre diferentes vertentes. A informação que resulta da audimetria, e, logo, baseada em métodos automáticos e menos dependentes do erro humano e segundo processos de amostragem probabilísticos, contínuos e permanentes, apresenta uma maior validade e fiabilidade.

O registo audiovisual, inclusive, apresenta grandes desvantagens, sendo considerado como um método que não respeita os direitos humanos. Apesar de, na investigação qualitativa se estar a aderir a estes processos, as suas limitações são evidentes nomeadamente, apesar de permitir inquirir o entrevistado no seu lar, não interfere em menor grau nos resultados alcançados. O telespectador poderá sentir-se intimidado, e por outro lado, é difícil extrair informação do material gravado.

As limitações que a audimetria apresenta também não ficam resolvidas com os métodos alternativos de medição das audiências. Além do mais, não é só na indústria televisiva que têm sido adoptadas tecnologias que permitem caracterizar os padrões de consumo da população, por exemplo, a análise de bases de dados transaccionais permite aos hipermercados a definição das estratégias de Marketing mais apropriadas e a caracterização das opções de compra da população. Na área industrial, e ao nível da contabilidade de gestão, as tecnologias ao nível informático são implementadas para facilitar a imputação de custos indirectos aos diferentes objectos de custeio.

A nível mundial, é reconhecido o contributo dos “*ratings*” para o meio televisivo e de radiodifusão. Os “*ratings*” permitem às empresas proporcionar um melhor serviço ao seu público e simultaneamente avaliar o seu grau de competitividade e sustentabilidade. Também contribuem para uma utilização mais eficiente dos *Media* e do controlo dos seus custos. Como vantagens da audimetria devemos salientar os seguintes aspectos:

(1) Certos comportamentos televisivos são analisados mais facilmente, como, por exemplo, o “*zapping*”.

Com um registo, ao segundo, da exposição televisiva podemos analisar determinados comportamentos para a totalidade do painel e no decorrer de todo o dia, e ano, o que de outro modo seria totalmente inviável; obtém-se um conjunto de informações relativas à evolução do consumo durante breves unidades temporais e que, consequentemente, permite medir a audiência de anúncios e de espaços comerciais e considerar a duplicação das audiências;

(2) Os resultados proporcionados pelo painel de audimetria podem ser analisados, logo após a sua detecção, uma vez que no dia seguinte à sua recolha já se encontram disponíveis. Recorrendo aos restantes métodos de medição das audiências, certamente, que a caracterização diária da exposição televisiva seria muito difícil e morosa, uma vez que são milhares os telespectadores que integram o painel;

(3) A audimetria permite resolver alguns problemas que surgem em áreas com elevados níveis de analfabetismo e favorece a incorporação de imigrantes na amostra. As variáveis sociodemográficas, entre outras, que caracterizam o painel influenciam a fiabilidade dos resultados recolhidos com outros instrumentos.

Um sistema de informação permite a recolha de dados mais fiáveis para os segmentos que apresentam uma menor formação académica, ou maiores e menores faixas etárias. Os dados de audimetria também são mais fiáveis porque não estão dependentes da capacidade de memorização do painel. Assim, contribui para uma menor probabilidade de alguém referir dados erróneos involuntariamente ou não, a respeito das suas selecções;

(4) A audimetria permite a recolha de dados de forma sistemática e continuada. Apesar de envolver elevados custos na sua implementação, apresenta certamente uma maior relação custo-benefício. Não será viável recolher a informação necessária ao planeamento de *Media* através dos restantes instrumentos. Os restantes métodos de medição das audiências proporcionam ao sector dados mais incompletos, menos fiáveis e objectivos, e apenas para um reduzido horizonte temporal e maiores unidades temporais;

(5) A audimetria corresponde a um instrumento de diagnóstico, planificação, tarifação e de previsão. Viabiliza a realização de previsões e extrapolações a partir de situações modelo verificadas no passado ou recorrendo a abordagens de previsão alternativas, e que estão em análise na presente tese;

(6) Os anunciantes precisam de conhecer a duplicação de audiências entre os suportes, para definirem a distribuição de contactos nas suas campanhas publicitárias, o que apenas será possível recorrendo à audimetria. Os inquéritos incidem em intervalos de tempo de 15 minutos, o que é insuficiente para os anunciantes, além de serem demasiado caros para o painel ser interrogado ao longo de várias semanas.

Os trabalhos que não recorrem a um sistema de informação não permitem obter informação a respeito dos efeitos dos planos de *Media*; não permitem conhecer a distribuição de contactos entre o público objectivo e os anúncios e não proporcionam informação a respeito da duplicação de audiências entre dois anúncios, situados em diferentes suportes. Os painéis medem comportamentos diários de forma contínua e, conseqüentemente é viável o cálculo da distribuição dos contactos históricos de uma campanha.

Como desvantagens da audimetria salienta-se, por exemplo, que poderá influenciar os comportamentos dos telespectadores. Uma vez que quem integra o painel sabe que está a ser vigiado, este facto poderá condicionar e influenciar o seu comportamento, assim como sucederia em maior grau se existisse um registo audiovisual do seu comportamento; as entrevistas também influenciam as respostas obtidas. Com base na mera interrogação e, sem controlo externo, também não será exequível recolher informação fiável, uma vez que as respostas estarão influenciadas pelas falhas de memorização e registo humano, assim como por uma alteração intencional dos resultados por forma a reflectir um comportamento que seja considerado mais correcto e valorizado por terceiros.

Com a audimetria, sendo este um processo que apenas depende do correcto manuseamento do telecomando, os telespectadores, por vezes, até se podem “esquecer” que os seus movimentos estão a ser registados, contribuindo para uma maior proximidade entre os padrões identificados e os que resultariam em contextos reais. Dificilmente os telespectadores, constantemente, e de uma forma sistemática, ao longo do ano terão bem presente que as suas escolhas estão a ser registadas e alterar as suas decisões que decorrem da sua vontade.

Segundo *Beville (1988)*, os *ratings* também apresentam algumas desvantagens entre as quais as seguintes, apresentando-se algumas considerações a seu respeito:

(1) Os *ratings* são estimativas e não medidas exactas da dimensão da audiência, devendo-se ter em conta as suas limitações metodológicas, como por exemplo, as amostras são por vezes pequenas, as taxas de cooperação são reduzidas e certos grupos étnicos encontram-se mal representados.

Devemos referir que, apenas um senso total garante a total representatividade da amostra mas, perante a sua impossibilidade, são utilizadas as sondagens parciais, nas quais se baseiam a totalidade dos métodos de medição de audiências. Portanto, toda a sondagem corresponde a uma estimativa, existindo erros relativamente à amostra. Por conseguinte, quanto menores forem os erros, maior qualidade terá a sondagem e a audimetria promove a sua minimização.

Os componentes da amostra exercem uma grande influência nos resultados finais, logo, as respostas incorrectas provocam a deturpação do tratamento estatístico. Toda a sondagem deve ser fiável e válida, sendo estes os parâmetros que determina a sua qualidade (*Leon, 1990*). A validade e fiabilidade da audimetria são demonstradas no *Apêndice A.2.5*.

Relativamente à dimensão da amostra, os valores adoptados a nível nacional enquadram-se nos parâmetros internacionais que referem que a dimensão mínima deverá ser de 300 lares (recomendação da UER). Em finais de 1997, a União Europeia de radiodifusão, que agrupa sensivelmente cem canais de televisão e de radio, publicou um relatório a respeito das técnicas de medição das audiências a nível Europeu. A uma dimensão de 300 lares corresponde, segundo *Dominick, Sherman e Copeland (1990)*, um erro de 7% na amostra e fornece um intervalo de confiança de 99%. Em Portugal, para o período em análise, a amostra corresponde a 850 lares e 1000 lares. As amostras devem ser representativas, ou seja, as respostas devem ser generalizáveis, com rigor científico, e devem proporcionar um retrato preciso da população a investigar, permitindo realizar deduções com exactidão (*Tille, 1990*).

Quanto às taxas de cooperação, na literatura da especialidade, está bem documentada a crescente aversão do painel para conceder entrevistas e o curto espaço de tempo após o qual começa a existir sinais de cansaço, assim como a dificuldade em obter dados fiáveis para determinados segmentos, recorrendo a outras vias que não implique a utilização das tecnologias de informação. Uma das vantagens da audimetria corresponde, precisamente, a promover, por exemplo, a integração de imigrantes e da totalidade dos grupos étnicos no processo de sondagem.

No que concerne a representatividade das amostras, as variáveis de segmentação utilizadas a nível nacional, correspondem às que são utilizadas também na maioria dos sistemas de audimetria internacionais, assim como em qualquer processo de sondagem, cujo objectivo seja investigar determinado aspecto nos comportamentos da população. Logo, é inevitável a sua aplicação no contexto da medição das audiências. As classes sociais *marktest* são um modelo de categorização social, que procura reflectir a capacidade económica dos consumidores e, não, criar um modelo de utilização universal. Apenas as amostras seleccionadas através de métodos probabilísticos, como se verifica com a audimetria, permitem estabelecer inferências precisas a respeito do resto da população, já que estas técnicas se baseiam no pressuposto de que todos os que integram o universo podem ser seleccionados como participantes da amostra.

2) Os *ratings* medem a exposição aos programas, canais e estações televisivas e, conseqüentemente, não reflectem necessariamente a satisfação do público. Os números apenas reflectem o que foi seleccionado e não traduzem a real qualidade dos programas.

Obviamente que estes dados apenas quantificam a exposição televisiva, permitindo a identificação dos programas que são seleccionados pelos telespectadores. A selecção poderá corresponder à alternativa que é considerada como sendo a melhor nesse momento, ou resultar das inúmeras influências estruturais, ou corresponder a programas que o público realmente considere como sendo mais interessantes.

(3) Os *ratings* não são bem utilizados frequentemente, uma vez que, conduzem ao cancelamento de programas de qualidade e entorpecem a criatividade. As decisões ao nível da publicidade e da estruturação da grelha de programas, são tomadas em função dos níveis de exposição ao painel e em detrimento da qualidade dos conteúdos, com o intuito de manter as empresas financeiramente saudáveis, e assim, subsidiar certo tipo de programas que não são auto-suficientes.

Certamente, que a estruturação de uma grelha de programação encontra-se fortemente condicionada pelos seus níveis de exposição. Contudo, o Canal 2, que pretende exercer uma função de serviço público promovendo a transmissão de programas considerados de maior qualidade, apresenta valores irrisórios de *ratings*. Um fraco desempenho, quantificado através dos valores de *ratings*, seria totalmente inviável para um canal privado que se encontra dependente das receitas publicitárias. Conseqüentemente, uma grelha que é considerada como apresentando uma “maior qualidade”, poderá ser seleccionada por um reduzido número de telespectadores. A nível nacional, uma elevada percentagem do painel, apresenta uma maior exposição a programas de ficção ou de entretenimento, e a tipologia de informação acaba por ficar em maior grau circunscrita apenas aos principais espaços informativos.

No texto de *Souchon (2003)*, consta uma retrospectiva da medição de audiências, além de serem mencionadas as principais linhas de força da actividade aplicadas à realidade francesa.

O autor salienta a importância da audimetria e pretende mostrar a importância do uso inteligente e fecundo do instrumento audimétrico:

Diz a velha sabedoria dos pedagogos: para ensinar matemática a João, é preciso conhecer-se a matemática e João. Para fazer uma boa política de programas, é preciso conhecer-se o público na sua diversidade de expectativas e comportamentos. Ora, a audimetria ajuda, do mesmo modo que os trabalhos económicos e sociológicos, assim como os trabalhos de prospectiva. (p.159)

A tecnologia e os sistemas de informação são cada vez mais utilizados e, têm sido sucessivamente aperfeiçoados, para promover a extracção de conhecimento útil e o bem-estar da sociedade. Os métodos qualitativos não são os únicos que podem ser considerados como sendo mais adequados para a análise do consumo televisivo. As escolhas metodológicas devem ser feitas em função dos recursos disponíveis, e da tipologia de dados que permite responder a determinadas questões. Qualquer escolha realizada, acaba por excluir outras possibilidades, e nenhum método deverá garantir ou deitar a perder determinado estudo, nem reclamar uma maior virtude sobre os restantes.

Assim sendo, uma vez que a questão em análise define as fontes que devem ser adoptadas, quando está em causa a previsão das audiências, não podemos deixar de recorrer a uma tecnologia que disponibilize um conjunto de dados com as características que apresentam as bases de audimetria, ainda que, complementada com fontes alternativas de dados em determinados contextos. Um painel corresponde à metodologia idónea para medir evoluções de comportamentos. Através da realização de inquéritos, o que é declarado não cumpre os requisitos do que é quantificado, medido, onde deve existir objectividade e ausência de estímulos internos e externos ao próprio indivíduo. Um inquérito também não representa um instrumento de previsão. Consequentemente, na presente tese privilegia-se a audimetria uma vez que corresponde, também, a um instrumento que proporciona a informação necessária à estruturação da grelha de programação e ao investimento publicitário. Será, inevitavelmente, o desenvolvimento dos sistemas de informação que, certamente, irá permitir que os melhores resultados sejam alcançados no futuro, ao nível do planeamento de *Media*.

A.1.5. O audio-matching

A CAEM, que representa a Comissão de Análise de Estudos de Meios, convidou as empresas Marktest, Kantar Media/TNS, GFK, AGB Nielsen e Euroexpansão para o concurso do novo sistema de audimetria realizado em 2011, enviando cartas com informação sobre o caderno de encargos (especificações técnicas e projecto de contrato). A empresa Marktest e a empresa GFK foram as finalistas do concurso, mas a GFK foi a vencedora.

Inicialmente, os canais de televisão defenderam a necessidade de existir um novo rumo com a multinacional GFK, ao passo que a maioria dos anunciantes e das agências de meios consideravam que a Marktest apresentara uma proposta mais sólida e conhece bem a realidade nacional, tendo apenas de adaptar o serviço às inovações actuais. Porém, posteriormente, as estações televisivas também apresentaram algum descontentamento face aos dados facultados. A tecnologia utilizada pela GFK designa-se de “*audio-matching*” (Fig. A.1.5.1) e corresponde a um método de amostragem de som para a detecção dos canais, assegurando assim, uma rigorosa medição ao segundo, do espectro analógico, digital e por cabo. A tecnologia de *audio-matching* capta o som emitido pela TV que é, posteriormente, comparado com o som recolhido na estação de captura da empresa, permitindo dessa forma a identificação dos canais e programas a que os telespectadores assistem em cada minuto.

Como vantagens desta tecnologia podemos salientar que apresenta independência das emissoras, o que nos garante a total independência dos seus resultados, e a facilidade de implementação nos lares.

A par do que se verifica na audimetria, esta tecnologia também permite a medição do consumo ao segundo e em “*time-shift*”, até sete dias após a transmissão dos programas. Esta tecnologia, não depende do tipo de plataforma de distribuição que é utilizada (analógica, digital, via cabo, entre outras), isto é, em princípio, qualquer plataforma pode ser medida com o *audio-matching*.

De acordo com o *site* institucional da empresa Marktest Audimetria, também se verificou a progressiva integração de alguns lares com a tecnologia de *audio-matching* em 2011, tendo a primeira fase ocorrido em Junho com 120 lares; a segunda em Agosto, com 80 lares e a última terminou no dia 27 de Setembro, com 40 lares.

Estes *meters*, desenvolvidos pela *kantar*, estão a ser utilizados em painéis de vários países, nomeadamente em Espanha e no Reino Unido, tendo sido, neste último, aprovado e recomendado pelo *BARB - Broadcasters' Audience Measurement Board* como o sistema oficial de medição de audiências de televisão. A inclusão destes audímetros permitiu fazer face a determinados problemas tecnológicos, para os quais, desde há longa data, que a Marktest Audimetria tinha vindo a alertar o mercado.

A.2. A DESCRIÇÃO DOS INSTRUMENTOS E PROCEDIMENTOS DE RECOLHA E ESTRUTURAÇÃO DOS DADOS UTILIZADOS

A.2.1. A descrição do processo de audimetria

Neste apêndice descreve-se sucintamente a tecnologia de audimetria, que corresponde à fonte de dados adoptada, a constituição do painel, método de amostragem e as bases de dados que foram disponibilizadas para a realização das análises propostas neste trabalho.

A audimetria é um dos processos que permite medir as audiências televisivas com o maior rigor e fiabilidade, proporcionando informação com uma elevada qualidade estatística. Sendo a área da electrónica um meio muito dinâmico, temos assistido a nível internacional a uma considerável evolução no âmbito dos métodos de medição das audiências (*Huertas, 2002*).

Há a salientar nomeadamente, a tecnologia de audimetria activa, a audimetria passiva e a PPM's, que representa a audimetria portátil. A audimetria possibilita a medição diária das audiências televisivas através de uma rede de audímetros, distribuída por todo o país.

Nas análises realizadas neste trabalho, recorreu-se a bases de dados que resultam de um processo de audimetria activa, na qual, um sistema centrado em um painel de lares regista a exposição individual através da informação que é transmitida pelos telespectadores. Nos lares que integram o painel é instalado um audímetro por televisor e é atribuído um número no telecomando a cada indivíduo com mais de 4 anos, pertencente ao agregado familiar. Sempre que a televisão estiver ligada, cada indivíduo deverá assinalar a sua presença através do telecomando e, logo, designa-se de audimetria activa.

O audímetro capta informação sobre o que está a ser visionado e regista os instantes em que um telespectador liga e desliga o televisor ou realiza uma mudança de canal, ou seja, apenas regista as horas e o canal a que este assiste. Esta informação, designada por dados dos ficheiros base de audimetria, é enviada via telefone para a Marktest Audimetria, onde será processada, validada e ponderada, permitindo o cálculo das audiências televisivas. A formatação torna possível formar a base de dados de audiências, de um dia, de audimetria. A validação, consiste na aplicação de uma série de normas e critérios aos dados, que permite eliminar e corrigir os comportamentos que não se enquadram em determinados critérios. A ponderação é um processo iterativo, através do qual se atribui um peso a todos os lares e indivíduos produzidos para um determinado dia, e a sua soma deverá representar o universo.

Do processo anteriormente descrito, resultam duas bases de dados, uma das quais associada a índices de audiências de canais, programas e espaços publicitários (*Software Telereport*) e outra, que permite a análise de anúncios, inserção a inserção, e das campanhas publicitárias (*Software Publitec*). A partir dos dados dos ficheiros, base de audimetria, calculam-se os índices de audiências para cada segmento sociodemográfico representado nas bases de dados, assim como para o universo. Correspondem a determinados agregados, ou seja, a taxas de audiência para determinado conjunto de telespectadores, como, por exemplo, as taxas de audiências para o sexo feminino, ou para a região de Lisboa, ou para a classe social D.

Os audímetros consistem em aparelhos electrónicos que possibilitam monitorizar várias fontes de sinal (televisores, vídeos, receptores de satélite, etc.) através da detecção do ON/ OFF, entradas auxiliares, reprodução e gravação e canal sintonizado. O processo é auditado por empresas especialistas e, os estudos coincidentais, são realizados para se demonstrar o grau de concordância que existe entre os dados recolhidos pontualmente através de entrevistas telefónicas e os dados que resultam do processo de audimetria.

A.2.2. A descrição do painel de audimetria

As características dos sistemas de informação, que permitem a recolha dos dados necessários ao cálculo dos índices de audiências, variam de país para país, em função das necessidades do mercado e dos recursos disponíveis. O processo de audimetria nacional, certamente que poderia ter sido melhorado em determinados aspectos, inculcando-lhe funcionalidades que promoveriam uma melhor caracterização qualitativa da interacção da audiência com o *Media*, tendo sido desenvolvidos esforços a diversos níveis neste âmbito, mas, obviamente, que todo o investimento será feito em função da sua viabilidade económica, para além de outros factores. Em 2012, esta tecnologia foi substituída pelo *audio-matching*, que a Marktest Audimetria também tinha adoptado parcialmente antes da realização do último concurso que visou a adjudicação do serviço de medição das audiências televisivas.

Para o período em análise, o painel associado à tecnologia de audimetria activo e individual corresponde a 850 lares, dos quais apenas 250 têm TV Cabo. As suas variáveis de segmentação são as regiões geográficas, as faixas etárias, as classes sociais, o sexo, a existência de TV Cabo e ser responsável pelas compras do lar (*Quadro A.2.2.1*).

A representatividade do painel é assegurada através de censos do INE e de estudos realizados pela Marktest e denominados como “*Establishment Survey*”, os quais têm como intuito caracterizar a população relativamente a uma série de variáveis sociodemográficas, ficando, assim, determinada a amostra teórica de lares que abrange apenas Portugal Continental.

Geográficas	Faixas Etárias
Grande Lisboa	Faixa Etária 4/14
Grande Porto	Faixa Etária 15/24
Litoral Norte	Faixa Etária 25/34
Litoral Centro	Faixa Etária 35/44
Interior	Faixa Etária 45/54
Sul	Faixa Etária 55/64
	Faixa Etária +64
Grupos Ocupacionais / Classes Sociais	Situação no Lar
Classe Social A/B - Alta/Média Alta,	Dona de Casa
Classe Social C1- Média,	Não Dona de Casa
Classe Social C2 - Média Baixa	(Responsável, ou não pelas compras do lar)
Classe Social D – Baixa	

Como processo de amostragem, recorrem à amostragem estratificada desproporcional. Resumidamente, no método de amostragem desproporcional, a dimensão de cada estrato na amostra não é proporcional à sua dimensão na população, mas, sim, desproporcional, uma vez que a sua dimensão está condicionada pela variabilidade observada nos dados.

As “*classes sociais marktest*” são um modelo de categorização social que foi definido por uma equipa multidisciplinar e que procura representar a capacidade económica dos consumidores.

Quadro A.2.2.1 - Caracterização das variáveis de segmentação do painel.
Fonte: Site institucional da Marktest.

Segundo *Fernandes (2001)*, as bases de dados de audimetria apresentam alguns erros na construção das diferentes classes sociais que servem de base à caracterização do perfil do público em especial, no que se refere ao indicador rendimento mensal familiar. Contudo, não levanta objecções à terminologia adoptada.

Em *Soares (1996)* também se aborda esta temática, que constitui uma dissertação de licenciatura em Sociologia. Em Marketing, a estratificação social dos consumidores, distingue-se das restantes variáveis pela existência implícita, ou explícita, de pressupostos teóricos e pela sua dependência face aos períodos temporais, e portanto, de determinadas actualizações periódicas.

A este respeito, podemos referir, que durante o ano de 1992, uma equipa multidisciplinar da Marktest realizou uma investigação pormenorizada sobre esta problemática. Deste estudo, resultou uma categorização que se encontra fundamentada estatisticamente, e que foi utilizada de 1993 a 1998. Esta categorização tinha como indicador subjacente o rendimento mensal líquido do agregado familiar (“RMLF”) no qual o indivíduo está inserido. No entanto, devido às limitações que surgem quando pretendemos conhecer o rendimento dos consumidores (principalmente ao nível das abstenções), não se utilizou o RMLF, mas sim, um conjunto de sete variáveis sem as referidas limitações e que, têm simultaneamente a capacidade de reflectir a posição de um indivíduo perante aquela variável. As variáveis utilizadas foram “a Situação na Ocupação”, “Ocupação”, “Instrução Escolar”, “Sexo e idade”, “Composição do agregado familiar” e “Regiões Marktest. Do modelo então criado, resultaram cinco grupos com uma distribuição percentual que se descreve no *Quadro A.2.2.2*.

Classe Social	Dimensão		
	1992	1998	2003
A	4,5%	4,9%	5,5%
B	10,7%	10,5%	11,9%
C1	25,3%	27,5%	24,9%
C2	27,6%	29,8%	31,0%

Em 1992, realizou-se uma segunda fase de actualização e aperfeiçoamento deste modelo a partir da experiência que foi alcançada, assim como de alguns aspectos que na altura se considerava ser necessário aprofundar.

Quadro A.2.2.2 – As classes sociais Marktest

Fonte: Site institucional da Marktest

A partir da matriz de cruzamento das variáveis “grupos ocupacionais” e “instrução” (*Quadro A.2.2.3*), preenchida com a população presente em cada célula, foram-se agrupando células por processo cumulativo, tendo-se definido o limite de cada grupo, ao atingir dimensões próximas (tanto quanto possível) da respectiva classe social no sistema criado na fase de 1992. Este processo de investigação, que culminou com a quantificação das classes sociais Marktest apresentada em 1998, teve por base os Censos realizados em 1991 pelo INE.

Os Modelos de Exposição Necessários à Aquisição de Publicidade no Sector Televisivo

NÍVEL	ANALF/ PRIMÁ RIA INCOM P	PRIMÁ RIA COMPL .	CICLO PREPA.	9º AÑO	10º/11º/ 12º	CURSO PROFIS / ARTIST	CURSO MEDIO/ FREC UNIV/B ACH	NOVAS LICENT	LICENC.	MASTER /PG/ DOCTOR.
OCUP.										
Pens.Dese np.Viv.Ren d, e Dom.Ref	D	D	D	D	C2	C2	C2	C1	C1	C1
R.trab E/Q e N E/Q	D	D	D	C2	C2	C2	C1	C1	C1	C1
Trab n E/Q	D	D	C2	C2	C2	C1	C1	C1	C1	C1
Trab E/Q	D	C2	C2	C2	C1	C1	C1	C1	C1	C1
R.Serv/ Com /Adm	C2	C2	C2	C1	C1	C1	C1	C1	C1	A
R. Téc.Esp / Peq.Prop.	C2	C2	C1	C1	C1	C1	C1	C1	B	B
Peq.Prop.	C2	C1	C1	C1	C1	C1	C1	C1	B	B
Emp.Serv/ Com/Adm	C1	C1	C1	C1	C1	C1	C1	B	B	B
Técnicos Espec.	C1	C1	C1	C1	C1	B	B	B	B	B
R.Q.Méd /Superiores	C1	C1	C1	C1	B	B	B	B	B	B
Quadros Médios	C1	C1	C1	B	B	B	B	B	B	A
Quadros Superiores	C1	C1	B	B	B	B	B	B	A	A

Quadro A.2.2.3 - Caracterização das classes sociais (Ocupação / Nível de Estudos)

Fonte: Agência de meios Tempo-Omd

Em 2003, tendo por base os censos de 2001, as classes sociais marktest foram redimensionadas para os valores obtidos nesta actualização. O cálculo das classes sociais ao nível do concelho foi realizado com recurso a uma extensa base de dados proveniente de entrevistas realizadas nos anos de 1999 e 2000 pela Marktest, nos diversos estudos regulares efetuados por esta empresa.

Em suma, as classes sociais Marktest são definidas com base no cruzamento de grupos ocupacionais e grau de instrução, uma vez que não se dispõe de informação fidedigna quanto ao rendimento económico dos telespectadores. *Huertas (2002)* declara que deve existir um modelo unânime destinado ao sector, sendo que a classe social encontra-se definida de acordo com as tarefas laborais, estudos e do rendimento:

Conviene remarcar que la clase social ha sido una de las variables que más problemas ha planteado y está planteando en este terreno. Con la implantación de la sociedad industrial, la clase social se constituyó en uno de los principales baremos a la hora de diferenciar a una persona de otra y de determinar su posición en un orden jerárquico social. El desarrollo de los sistemas de producción de producción fabril creó diferentes categorías de trabajo, ... , de manera que las tareas laborales servían para establecer la clase social a la que pertenecía. La evolución del capitalismo ha seguido perfilando esta cuestión, y la cantidad y la calidad de bienes adquiridos por una persona también se han convertido en indicadores de su estatus. ... La práctica más consolidada es que haya una única medición en cada país, para que los usuarios de las mismas, los sectores televisivo y publicitario, no dispongan de datos contradictorios a la hora de firmar sus acuerdos comerciales. Esto supone que ambas partes deben estar de acuerdo en el modo en que se opera el sondeo, en toda su extensión ... Que variables son las más adecuadas para definir la clase social a la que pertenece cada uno de los telespectadores? Las de uso más extendido son: el nivel de estudios/profesión y la posesión de bienes, o el cruce de ambas, siendo lo más habitual la última opción. (p.31)

Consequentemente, segundo a bibliografia da especialidade, a forma de construção das classes sociais Marktest corresponde aos parâmetros adoptados internacionalmente, recorrendo ao grau de instrução e aos grupos ocupacionais que se definem com base nos seus índices económicos. Em *Huertas (2002)*, também se questiona a relevância das variáveis sociodemográficas e da classe social quando se pretende a análise da exposição televisiva:

Es la clase social determinante del consumo mediático? Toda vez que el uso de la televisión está tan extendido, tiene sentido desglosarse la audiencia por clases sociales?... El consumo televisivo está tan generalizado que apenas se dan diferencias entre los diversos grupos sociales, establecidos éstos por edad, género, lugar de residencia, estado civil o clase social. En nuestro país, encontramos una mínima distinción entre los grupos mas jóvenes y los de mayor edad y también se da un escaso margen de diferencia entre aquellas personas con diferente grado de estudios. (p.32)

Nas análises propostas as variáveis sociodemográficas foram utilizadas com o objectivo de avaliar a sua repercussão na evolução dos “*ratings*”. Também devemos referir que, em Sociologia, não existe consenso sobre a noção de classe social; na norte-americana, especialmente, este conceito tendeu, inclusive, a ser esvaziado do seu significado original, diluindo-se e confundindo-se com o de estratificação social (*Stavenhagen, 1973*).

Em *Reyes (2009)*, são apresentadas algumas considerações a este respeito. O autor salienta, que as noções de estrato e de classe social, são utilizadas por vezes como sinónimos e, outras vezes, em contextos diferentes, e inclusive antagónicos, de concepção da desigualdade por estarem associados a abordagens teóricas divergentes, como, por exemplo, ao Funcionalismo e ao Marxismo. Também, podem designar conceitos de diferente grau de generalização, referindo-se a classe social a um sistema ou variante particular de estratificação. O autor refere que, à margem dos problemas de terminologia, existe uma grande controvérsia a respeito da natureza, tipos e elucidação da estratificação social. Em primeiro lugar, um contexto de debate segue as linhas da polémica nominalismo / realismo. Um segundo aspecto, corresponde ao carácter unidimensional ou multidimensional da estratificação, para além da controvérsia a respeito de, se a estratificação equivale a uma realidade universal e inevitável e inerente a toda a sociedade humana, ou se pelo contrário, é um fenómeno contingente e erradicável. O mesmo autor salienta que o debate contemporâneo centrou-se no confronto que existe entre a teoria Funcionalista e as teorias Conflictivistas.

Para o sociólogo Anthony Giddens, o conceito de classes ainda está longe de ser definido (*Bensi, 2009*). Cada corrente ideológica possui as suas definições, e a academia não tem uma posição clara para o termo, tendo em vista este conceito ser compreendido e utilizado de diversas formas no quotidiano. Por exemplo, os conceitos de camadas sociais ou de estratos sociais foram utilizados como sinónimos de classe por sociólogos Funcionalistas, que examinaram em profundidade as correlações existentes e inexistentes entre o prestígio social e a riqueza, ocupação, educação, raça e religião, em particular. Consideram a classe social como uma questão de estratificação social que, de acordo com *Parsons (1940)*, é uma classificação hierárquica dos indivíduos de um sistema social.

Um segundo aspecto está relacionado com a referida heterogeneidade do consumo televisivo, que nem sempre é tão significativa quando se avalia a repercussão das características sociodemográficas, sendo necessário identificar quais são os elementos que perfilam a diversidade. Com base nas análises realizadas, também foi possível constatar a proximidade de valores para determinados níveis das variáveis de segmentação, identificando-se os macrosegmentos existentes.

Na opinião de *Huertas (2002)*, entre as variáveis sociodemográficas que tradicionalmente se utilizam, a faixa etária será a mais explorada e menos conflituosa. Releva que, as diferenças são muito subtis, e apenas grandes segmentos de telespectadores, podem apresentar uma maior diferenciação ao nível da exposição televisiva. As fronteiras entre os segmentos são difusas e apenas temporárias, isto é, as pessoas podem pertencer simultaneamente a vários grupos e podem movimentar-se entre eles.

Na presente tese, as variáveis sociodemográficas são utilizadas para diferenciar os vários segmentos quanto à respectiva evolução dos indicadores de audimetria, uma vez que a previsão dos *ratings*, pode resultar das análises realizada para segmentos de telespectadores. Do ponto de vista das necessidades de informação que a previsão das audiências poderá colocar ao sector, pretendemos verificar qual é a segmentação do painel em função da variabilidade que incutem aos principais indicadores de audimetria, uma vez que se esperam diferenças subtis entre um grande número de níveis associados às diversas variáveis de segmentação. Poderá ser necessário prever os valores de alguns indicadores para o universo televisivo, à custa dos seus valores para determinados segmentos mais homogêneos relativamente à evolução das variáveis.

A.2.3. O método de amostragem

A Marktest Audimetria utiliza o método estratificado desproporcional para definir a amostra, quer dizer, a amostra é seleccionada em várias etapas, segundo o método aleatório estratificado por regiões com tamanhos não proporcionais.

Frequentemente, considera-se que a dimensão populacional de cada estrato deve determinar a dimensão da respectiva subamostra, ou seja, a amostra total deverá ser distribuída pelos estratos na proporção directa do tamanho de cada estrato.

No entanto, a precisão de uma amostra, não depende apenas da dimensão da população mas também da respectiva variabilidade. A variabilidade de um estrato é elevada, quando os seus elementos apresentam características muito heterogéneas. Esta situação implica que um estrato com uma maior variância, poderá levar à selecção de um maior número de unidades amostrais, quando comparado com um estrato com a mesma dimensão populacional, mas com uma menor variância.

Este método optimiza a amostra aplicada a um universo estratificado, pelo que também é, frequentemente, apelidado de distribuição estratificada optimizada. A nível geográfico, por exemplo, os estratos mais urbanos costumam apresentar comportamentos mais heterogéneos do que os estratos com um maior índice de ruralidade, pelo que, uma amostragem desproporcional permite obter dados mais rigorosos, através de uma sobre amostragem nas regiões que são mais urbanas. Se a população de um estrato apresentar uma elevada homogeneidade do consumo televisivo, será necessário um menor número de observações para considerar as diferentes situações que caracterizam essa mesma população.

A aplicação de uma amostragem desproporcional quando comparada com uma distribuição proporcional, reduz a margem de erro de cada estrato, mas implica que o cálculo conjunto de dois ou mais estratos, tenha que ser sujeito a uma calibragem (aplicação de ponderadores) na fase de tratamento dos dados. Dessa forma, cada estrato contribuirá para os resultados, segundo o peso real e não o peso amostral. Todos os estudos por amostragem apresentam desvios em relação aos valores reais para o universo. Esses desvios podem ser agregados em dois grandes grupos: o erro amostral e erro não amostral. O primeiro é mensurável e varia na ordem inversa da dimensão da amostra ou subamostra em análise. O segundo, não é mensurável, e advém de factores como perguntas mal desenhadas, entrevistas mal aplicadas ou erros de introdução dos dados.

A.2.4. A caracterização das bases de dados

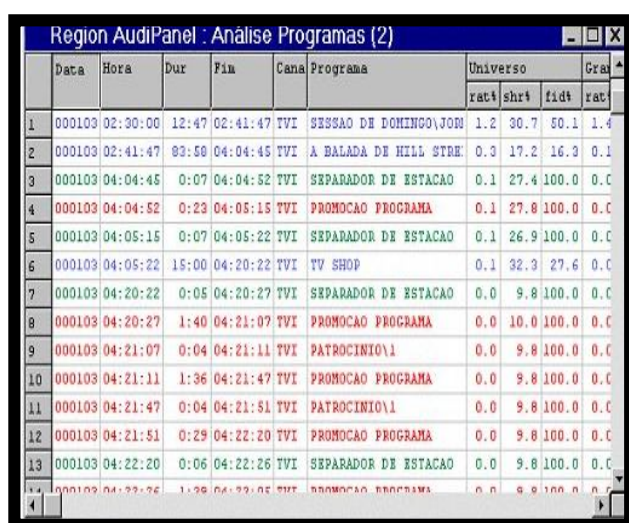
Através da audimetria o utilizador final poderá ter acesso a duas bases de dados, numa das quais constam os índices de audiência programáticos e de blocos de anúncios, designada de *Telereport* e uma segunda base de dados *Publitem* onde constam as audiências das inserções publicitárias.

A.2.4.1. A caracterização da base de dados *Telereport*

O *software Telereport* processa os dados das audiências televisivas, permitindo realizar as seguintes análises:

➤ Análise de Tabela de Programação

Nesta opção é elaborada uma tabela com a programação emitida nas datas indicadas e com alguma informação que a caracteriza: Identificação do Programa, Canal, Hora de Início e Hora de Conclusão, Duração do programa, Nível programático (global, parcial, intercalado) e a Tipologia programática, existindo três níveis.



	Data	Hora	Dur	Fin	Canal	Programa	Universe			Gra
							rat	shr	fid	
1	000103	02:30:00	12:47	02:41:47	TVI	SESSAO DE DOMINGO, JON	1.2	30.7	50.1	1.4
2	000103	02:41:47	03:50	04:04:45	TVI	A BALADA DE HILL STRE	0.3	17.2	16.3	0.1
3	000103	04:04:45	0:07	04:04:52	TVI	SEPARADOR DE ESTACAO	0.1	27.4	100.0	0.0
4	000103	04:04:52	0:23	04:05:15	TVI	PROMOCAO PROGRAMA	0.1	27.8	100.0	0.0
5	000103	04:05:15	0:07	04:05:22	TVI	SEPARADOR DE ESTACAO	0.1	26.9	100.0	0.0
6	000103	04:05:22	15:00	04:20:22	TVI	TV SHOP	0.1	32.3	27.6	0.0
7	000103	04:20:22	0:05	04:20:27	TVI	SEPARADOR DE ESTACAO	0.0	9.8	100.0	0.0
8	000103	04:20:27	1:40	04:21:07	TVI	PROMOCAO PROGRAMA	0.0	10.0	100.0	0.0
9	000103	04:21:07	0:04	04:21:11	TVI	PATROCINIO\1	0.0	9.8	100.0	0.0
10	000103	04:21:11	1:36	04:21:47	TVI	PROMOCAO PROGRAMA	0.0	9.8	100.0	0.0
11	000103	04:21:47	0:04	04:21:51	TVI	PATROCINIO\1	0.0	9.8	100.0	0.0
12	000103	04:21:51	0:29	04:22:20	TVI	PROMOCAO PROGRAMA	0.0	9.8	100.0	0.0
13	000103	04:22:20	0:06	04:22:26	TVI	SEPARADOR DE ESTACAO	0.0	9.8	100.0	0.0

Permite-nos também calcular os índices de audiências programáticos, a nível global e parcial, ou seja, considerando não só a duração total do programa como também diferenciando cada uma das suas partes. Ao programa global corresponde o nível 1, às partes programáticas o nível 2 e quando um programa é transmitido no decorrer de outro programa é identificado como pertencendo ao nível 3 (Fig. A.2.4.1.1).

Fig. A.2.4.1.1 – Tabela de programação do *software Telereport*
Fonte: Site institucional da Marktest

A identificação dos programas consta apenas para os quatro canais generalistas: RTP1, Canal 2, SIC e TVI. A informação é proporcionada para os segmentos, canais, variáveis e datas seleccionadas pelo utilizador. Não é possível obter através do *software* a informação referente apenas ao segundo nível programático, uma vez que os programas sem intervalo, encontram-se identificados como sendo do primeiro nível e são excluídos da amostra quando se selecciona o nível programático parcial em vez do global. Consequentemente foi necessário reestruturar estas bases de dados para serem realizadas as análises parciais.

Nestas bases de dados constam inúmeros índices de audiências, alguns dos quais são calculados a partir de outros. Os índices de audiência são estimados para a totalidade do painel (denominado de universo televisivo) e, também, para cada um dos segmentos representados nas bases de dados, por exemplo, para o sexo masculino vs. sexo feminino, região de Lisboa, região do Porto; classe social AB, etc.

No *software* Telereport constam ainda índices de audiências para blocos de anúncios, ou seja, desagrega os espaços comerciais em função da tipologia dos anúncios transmitidos (por exemplo, publicidade normal, patrocínios, promoções, separador de estação, etc.). Os índices de cada uma das inserções que integram determinado espaço comercial constam de uma outra base de dados, associada ao *software* Publitex. Estes blocos de anúncios são caracterizados com base nos mesmos índices que caracterizam os programas que lhe estão associados.

➤ Análise de Grelha

A análise de grelha constrói a curva diária de *Rat%* ou *Shr%* de cada canal e complementa-a com a respectiva grelha de programação. Desta forma, permite comparar graficamente a evolução do *rating* de cada canal no decorrer de um dia, ou comparar as curvas de *ratings* ou curvas de *shares* do mesmo canal em vários dias.

➤ Análise Minuto a Minuto

Estas análises permitem estudar a audiência dos períodos horários compreendidos no período de emissão, em qualquer intervalo de tempo: dias, semanas ou meses. Proporciona os índices de audiências por períodos, datas e segmentos-alvo seleccionados.

➤ Análises Individuais

Nas bases de índices de audiências programáticos também constam as denominadas análises individuais, as quais permitem realizar análises mais específicas a um evento ou conjunto de eventos, isto é, um programa, ou suporte horário, em determinado canal. Cada análise individual proporciona uma informação diferente relativamente à audiência de determinado evento.

➤ Análises Destino e Origem e de Duplicação

Através destas análises podem ser calculados, por exemplo, os índices de duplicação de audiências entre programas e os dados de fluxos de audiências, ou seja, a migração de audiências entre os vários eventos seleccionados. Os índices de duplicação representam taxas de audiências conjuntas entre dois programas, pertencentes ao mesmo dia, ou não. Correspondem a taxas de audiências, de certo modo, equivalentes a “probabilidades condicionadas”, quer dizer, para os programas seleccionados representam a percentagem da audiência presente no evento em coluna que também está presente nos eventos em linha.

Através do *software* devemos seleccionar, individualmente, cada um dos programas, ou partes de programas, que vão constituir a amostra sobre a qual se irá calcular o seu valor. Para a grelha de programação diária é necessário seleccionar cada um dos dias individualmente, uma vez que, o *software* não produz estes resultados, por exemplo, para uma semana. Através destas análises, é possível calcular uma estimativa dos denominados índices de lealdade programática, seleccionando os diferentes episódios de determinado programa transmitidos no decorrer do horizonte temporal estudado.

Os dados representam a percentagem de telespectadores que assistiram a um programa em determinado dia, e que também o fizeram num dia posterior. Também, nos permite efetuar o cálculo dos índices de duplicação entre programas adjacentes, transmitidos no mesmo dia e canal. Estes dados são proporcionados em quadros que cruzam os programas em análise, representando cada célula a percentagem da audiência presente no evento em coluna, que também está presente no evento em linha, por exemplo “dos telespectadores que assistiram ao espaço informativo da noite calcula quantos também assistiram à primeira telenovela da noite”. Nos dados de identificação da programação apenas se encontram representados a data, a hora, a duração, o canal e o nome do programa.

As análises “destino e origem” facultam a migração das audiências entre diferentes programas e canais, tais como, “Considerando um programa em concreto, calcula a percentagem da sua audiência que no decorrer dos programas antecessores e sucessores permanece no mesmo canal”; ou “assiste a pelo menos 50% da duração de um programa dos canais concorrentes”; assim como, a percentagem dos telespectadores que permanece com o televisor ligado ou que o desligam. As variáveis em análise correspondem a TV_ON, RTP1, RTP2, SIC, TVI, TV_Cabo, Satélite, Vídeo e TV_OFF e as análises são realizadas numa base diária.

➤ Exclusividade e Índice de Exclusividade

Para além das análises já mencionadas, também constam nestas bases de dados os índices de exclusividade, que conduzem a resultados simétricos aos das análises de duplicação das audiências. Ou seja, cada célula representa a percentagem da audiência do evento em coluna que não esteve presente no evento em linha. O índice de exclusividade resulta do quociente entre a percentagem de duplicação de um programa e a percentagem de exclusividade. Esta variável representa uma medida da cobertura exclusiva de um evento, em relação aos restantes eventos. Será superior a 100, quando a exclusividade do evento em coluna é maior do que a duplicação em relação ao evento em linha e menor que 100, quando a duplicação do evento em coluna com o evento em linha é maior que a exclusividade.

- Análise de Assiduidade – Análise da distribuição dos indivíduos de cada alvo por classes de percentagem de tempo visto da duração total dos eventos.
- Fidelidade – Média do número de contactos, GRP's, Cobertura total e distribuição de contactos total de um conjunto de eventos.

Nas bases de dados do *software* Telereport não consta explicitamente a contra-programação de cada programa, sendo necessário por isso, seleccionar cada um deles individualmente para ser possível a sua posterior identificação. Nas bases de audimetria a informação está excessivamente dispersa, não inter-relacionada e muito incompleta, sendo fundamental o cálculo ou a correcção de inúmeras variáveis, por forma a enriquecer as análises e para conseguir inter-relacionar a totalidade dos dados.

A.2.4.2. A caracterização da base de dados Publitex

Na base de dados associada ao *software* Publitex consta alguma informação relacionada com a transmissão das inserções publicitárias, assim como, os respectivos índices de audiências. Como por exemplo:

- Informação sobre as inserções publicitárias: marca, submarca do produto divulgado, sector, classe de produto divulgado, anunciante, agência, central de compra, localização e tipo de anúncio;
- Informação relacionada com a transmissão das inserções publicitárias: canal, código do programa, data, hora, duração, posição no espaço comercial, total de inserções no espaço comercial, tabela de preços associada e investimento;
- Ficheiro de *ratings*, onde consta a informação associada à transmissão das inserções publicitárias e para segmentos pré-definidos pelo utilizador.

Nesta base de dados encontram-se representados os mesmos segmentos sociodemográficos que foram mencionadas para a base de dados Telereport.

A.2.5. A validez e adequabilidade da fonte de dados

Um método de medição das audiências é válido quando mede o que se propõe medir. (Manyntz, Holm & Hubner, 1988). Goetz e LeCompte (1988) abordam dois tipos de validez: a validez interna e a externa.

Um método é válido internamente quando a análise dos resultados que viabiliza não origina conclusões que sejam incompatíveis. Um método apresenta validez externa quando os dados podem ser generalizáveis ao resto da população, ou seja, quando cumpre basicamente todos os requisitos que exige a representatividade, como é o caso dos métodos de amostragem probabilísticos. A audimetria cumpre claramente os dois requisitos de validez.

A fiabilidade de um método é verificada na medida em que a sua repetida utilização produz, sob as mesmas condições, idênticos resultados. A fiabilidade depende da sua precisão e exactidão. A validez sempre pressupõe a fiabilidade, mas para um estudo ser válido não é suficiente que seja fiável. Um método poderá ser fiável, mas se os indicadores escolhidos não forem adequados aos objectivos previstos, os seus resultados carecem de validez (Huertas, 1997). Assim como existem determinados procedimentos para detectar a validez, também existem para a avaliação da fiabilidade. Por exemplo, no âmbito da audimetria, são realizadas sondagens telefónicas, a uma sub amostra, para verificar os seus valores, e por vezes recorrendo a diferentes amostras de telespectadores.

Os métodos de medição de audiências podem ser classificados desde dois pontos de vista: a sua periodicidade e o tipo de amostra. Relativamente à periodicidade, estes diferenciam-se em contínuos ou pontuais. Os estudos contínuos podem ser periódicos ou permanentes, recolhendo estes últimos informações sobre a audiência de forma ininterrupta, como é o caso da audimetria. Os métodos contínuos por sua vez permitem estabelecer evoluções e detectar tendências, ao permitir comparar a audiência em diferentes períodos, sempre que as investigações sucessivas sigam idênticos parâmetros. Caso contrário não se podem confrontar os resultados.

As análises propostas na presente tese foram realizadas com base em amostras que correspondem ao mesmo horizonte temporal para os vários anos em análise, permitindo uma análise da evolução dos índices de audiências. Ao dispôr de uma análise diacrónica da audiência, permite estabelecer modelos mais precisos a respeito dos comportamentos e dos dados que deles resultam, o que vem reafirmar a utilização destas bases de dados como instrumentos de previsão.

Os instrumentos de medição que assumem um carácter pontual, como, por exemplo, os inquéritos, apenas permitem conhecer a audiência num período de tempo que corresponde ao do estudo, proporcionando uma visão estática das audiências. Isto impede a realização de previsões, uma vez que não reflectem uma evolução ou tendência. Este tipo de análises apenas se utiliza em situações excepcionais e quando os métodos contínuos são insuficientes, como acontece quando se transmite um programa que assume um carácter excepcional e especial.

Os métodos contínuos oferecem ao utilizador maiores garantias quando são realizadas análises de previsão (*Chaniac, 1988*) e do ponto de vista científico também apresentam algumas vantagens. Segundo *Durand (1982)*, as sondagens contínuas permitem ultrapassar o problema que representa a variabilidade das respostas dos elementos que integram as amostras.

Dependendo das características dos trabalhos poderá ser relevante, ou não, a sua periodicidade, porque a resposta poderá não variar dia após dia. Se uma pessoa pergunta a alguém durante 5 dias consecutivos que leite utiliza, provavelmente a resposta será sempre a mesma. No entanto, quando se interroga a população a respeito do programa a que assistiu na véspera, a resposta deve variar, em especial porque a exposição ao *Media* é irregular.

O audímetro regista as selecções que foram realizadas ao longo do tempo, informação que seria inviável obter para os 365 dias do ano e no decorrer de vários anos com base nos métodos alternativos. Por outro lado, o número de horas de emissão dos canais é cada vez mais alargado uma vez que, em alguns deles já não existem interrupções.

Isto torna os métodos contínuos no melhor instrumento uma vez que, podem detectar os comportamentos minuto a minuto. A construção de um painel, ou seja, de uma amostra fixa, resultou num grande avanço para o mercado televisivo, interessado em obter dados válidos e fiáveis ao instante e em geral no âmbito da sociologia empírica (*Durand, 1982*). As suas primeiras investigações centraram-se no consumo e na opinião pública.

Os métodos de amostragem podem ser probabilísticos, como no caso da audimetria, ou não probabilísticos. Apenas as amostras seleccionadas através de métodos probabilísticos permitem estabelecer inferências exactas a respeito do resto da população, uma vez que todos os integrantes do universo podem resultar seleccionados como membros da amostra. As análises probabilísticas são utilizadas na tecnologia de audimetria. Nestes métodos utilizam-se duas fontes de informação que nem sempre cumprem esses requisitos: os censos e as guias telefónicas. Consequentemente, a Marktest realiza estudos para avaliar a evolução das variáveis em estudo.

Dentro das técnicas probabilísticas existem as amostras aleatórias e as aleatórias estratificadas. No segundo conjunto identificamos dois procedimentos, o estratificado simples e proporcional ou desproporcional e o polietápico ou por agrupamentos.

Os métodos não probabilísticos, do ponto de vista económico são mais adequados devido ao seu custo ser menor. No entanto, do ponto de vista científico, são mais questionáveis, uma vez que nem todos os integrantes do universo apresentam a mesma probabilidade de vir a formar parte da amostra em questão, o que impede que sejam retiradas conclusões globais e extensíveis ao conjunto da população estudada. Estes sistemas surgem em países menos desenvolvidos do ponto de vista da comunicação social.

Nos métodos não probabilísticos, encontram-se as amostras acidentais, as amostras por quotas ou por grupos e as amostras incidentais ou de experta eleição:

- Amostras acidentais – escolha da amostra por inclusão de qualquer pessoa, por exemplo, as cem que o entrevistador encontra, como no caso dos inquéritos de rua;
- Amostras por quotas ou grupo – seleccionam-se as pessoas a partir de algumas indicações, por exemplo, de cada 3 homens seleccionar uma mulher. Esta técnica apresenta inúmeros problemas. Em primeiro lugar costumam ser pessoas do âmbito cultural do entrevistador e, frequentemente produz-se o efeito de bola de neve constituída por pessoas que são conhecidas de cada um dos interrogados;
- Amostras incidentais – um grupo de investigadores e de pessoas especializadas são encarregues de realizar o processo de extracção da amostra.

Devemos salientar que apenas com a utilização das amostras aleatórias é possível conhecer o grau de confiança que temos nos resultados das análises, ou seja, o grau de certeza que se tem a respeito da precisão da estimativa que foi calculada. Em contrapartida, são as amostras não aleatórias que viabilizam a extracção de conclusões da forma mais rápidas e com o custo menor (*Vicente, Reis & Ferrão, 1996*).

Quer se seleccione uma amostra aleatória, ou não, o importante será obter estimativas próximas dos parâmetros a estimar, e isto só se consegue, se tivermos uma amostra o mais representativa possível do universo em estudo. Por conseguinte, na área da audimetria, o método de amostragem seleccionado pela empresa Marktest corresponde ao que produz uma maior representatividade e precisão.

No *Quadro A.2.5.1* constam as principais diferenças entre os métodos probabilísticos e não probabilísticos de amostragem.

Os Modelos de Exposição Necessários à Aquisição de Publicidade no Sector Televisivo

Método / Descrição	Vantagens	Desvantagens
Métodos Probabilísticos		
Amostragem Simples Qualquer conjunto de n elementos tem a mesma probabilidade de ser seleccionado, de onde resulta que os elementos têm igual probabilidade de serem seleccionados)	Utilização fácil	Os membros de alguns grupos de interesse menos representativos podem não ocorrer nas proporções desejadas
A Amostragem Estratificada (população estudada é agrupada de acordo com características de interesse ou estratos)	Conduz a Análises por subgrupos com variâncias inferiores do que na amostragem simples	Caro e moroso quando existem muitos estratos
Amostragem Sistemática (todo o x-ésimo elemento da população é seleccionado até perfazer o tamanho da amostra, de acordo com um passo fixo. Esse passo é determinado dividindo o tamanho da amostra desejada)	Conveniente quando existe uma listagem de nomes como suporte da amostra	Dever-se-á ter em conta os padrões de repetição que podem enviesar a amostra
Amostragem por Cachos e Multi-etápica (Dos grupos formados naturalmente e que fazem parte da amostra serão inquiridos todo os seus elementos)	Utilização conveniente quando existem unidades estatísticas que correspondem aos grupos desejados (escolas, hospitais, etc.)	
Métodos Não Probabilísticos		
Amostragem por Conveniência (utilização de indivíduos que se encontram disponíveis)	Método prático pois a investigação recai em unidades já disponíveis (estudantes nas escolas, doentes na sala de espera, etc.)	Devido ao carácter oportunista da amostra, os seus elementos podem não ser representativos da população
Amostragem por “Bola de Neve” (Elementos previamente identificados identificam outros membros da população)	Útil quando não existem referências sobre a população ou essas referências são muito difíceis de obter	A amostra pode resultar bastante enviesada
Amostragem por Quotas	Torna-se prático quando existe informação fiável sobre as proporções dos atributos que interessam na população	Neste processo o entrevistador pode conferir involuntariamente enviesamentos na selecção dos inquiridos

Quadro A.2.5.1 - Vantagens e desvantagens dos métodos de amostragem probabilísticos e não probabilísticos. Fonte: ALEA - Acção Local Estatística Aplicada

APÊNDICE B

A INDÚSTRIA TELEVISIVA E A ACTIVIDADE DE PLANEAMENTO DE *MEDIA*

Neste apêndice, caracteriza-se o sector televisivo nacional e as suas perspectivas futuras, uma vez que a evolução dos indicadores de audimetria reflecte o desempenho das estações televisivas. Comparam-se as diversas estações ao nível dos conteúdos transmitidos e das estratégias de programação e contra-programação adoptadas. Posteriormente, consta a descrição da actividade de planeamento de *media*, referindo-se as diversas etapas que a integram e a informação necessária no decorrer da construção de uma campanha publicitária. Apresentam-se as questões que levantam maiores dificuldades e que são essenciais para a comercialização do tempo publicitário.

B.1. CARACTERIZAÇÃO DO SECTOR TELEVISIVO NACIONAL

A televisão, a nível nacional, surgiu em meados do século XX, por iniciativa do Estado. No decorrer do tempo, apresentou várias evoluções, como, por exemplo, o aparecimento dos canais privados generalistas e de sinal aberto e os canais por TV Cabo. De acordo com *Fernandes (2001)*, desde a década de 80, temos assistido a uma profunda transformação neste sector, e que resultou, quer do desenvolvimento tecnológico, quer das decisões económicas e políticas que conduziram a uma transformação da estrutura das televisões, a nível Europeu. A mesma autora, também, refere que esta fase tem sido considerada como a terceira era dos meios de comunicação, após o nascimento da rádio e da generalização da televisão, e que resulta do desenvolvimento das novas redes de telecomunicações e de uma progressiva introdução das tecnologias de informação.

Actualmente, existem quatro canais generalistas, dois dos quais são públicos e os restantes privados, e um mercado por subscrição que é difundido por satélite e por TV Cabo, que permite aceder a múltiplos canais nacionais e internacionais. Os canais privados vieram introduzir uma maior concorrência e dinamismo ao sector. A luta pela maior preferência dos telespectadores tornou-se no principal objectivo da maioria das estações televisivas, as quais divergem ao nível das estratégias de programação e contra-programação adoptadas e que, por sua vez, condicionam a evolução dos indicadores de audimetria.

B.1.1. As estações televisivas e a programação

Segundo *Fernandes (2001)*, com o início da actividade da televisão privada, em Portugal, a grelha de programas passou a constituir um instrumento de fidelização do público e de acção estratégica das estações televisivas. A autora refere que quando existia apenas o monopólio da televisão pública, os telespectadores eram vistos como um colectivo passivo e os programas eram definidos em função de três objectivos e que se traduzem em educar, informar e distrair. A televisão privada e o gradual acréscimo do número de canais transformou a oferta televisiva e a sua relação com o público, destacando-se, na actualidade, a vertente lúdica (*Fig. B.1.1.1*).

Actualmente, as televisões públicas e privadas, como consequência do sistema comercial competitivo que se instaurou entre as estações de televisão, apresentam como principal objectivo conseguir maximizar os seus *shares* de audiências. Deste modo, a audiência como público transformou-se numa audiência como mercado e esta transformação fez-se acompanhar de inúmeras tensões nas organizações de serviço público em termos de uma nova relação com a audiência (*Fernandes, 2001*). A lógica comercial domina a televisão portuguesa, na actualidade, em que a oferta é estruturada tendo em conta o papel da publicidade e do público. A televisão comercial depende da publicidade para sustentar a sua viabilidade, enquanto empresa de Comunicação Social. Consequentemente, surge o conceito de audiência como mercadoria e a programação televisiva é produto de estratégias que visam obter a audiência máxima.

No horário nobre a grelha de programas é estruturada de acordo com o perfil do público disponível e tendo em consideração o facto de toda a família estar reunida. As estratégias de programação são homogéneas e rotineiras, com o intuito de fidelizar o grande público. A repetição horizontal (“*stripping*”) representa um meio de fidelização, caracterizando-se pela transmissão do mesmo programa no mesmo horário de Segunda a Sexta e, também, se pratica a fidelização vertical, ou seja, que se mantém toda a semana.

No entanto, existe uma concepção de audiência televisiva diferente na televisão pública e na televisão privada: audiência como público e audiência como mercadoria, respectivamente.

Género	1999	2004	2009
Ficção	40%	35,9%	24,6%
Ficção Nacional	1%	16,2%	n.d.
Ficção Internacional	39%	19,7%	n.d.
Informação	7,5%	14,7%	15,4%
Entretenimento	4%	16,3%	22,8%
Diversos	6,7%	3,4%	1,5%
Juventura	16,8%	3,2%	8,3%
Cultura			
Geral/Conhecimento	2,1%	1,9%	1,1%
Desporto	3,4%	1,8%	1,1%
Publicidade	19.5%	22.7%	25.4%

Quadro B.1.1.1 – A evolução da programação

Fonte: Adaptado de CRECS (2006) e ERC (2010)

Isto fundamenta-se no facto da televisão pública e da televisão privada apresentarem missões diferentes, ou seja, à primeira compete o exercício do serviço público regulado pelo contrato de concessão e que implica uma responsabilidade social e cultural para com o público e a segunda assume uma vertente comercial dirigida para o índice de audiência mais elevado e que cativa os anunciantes.

O canal RTP1, e em especial os canais privados, SIC e TVI, apresentam uma grelha de programas denominada de federativa e que visa maximizar as audiências no decorrer da totalidade da emissão diária. De acordo com *Fernandes (2001)*, cada faixa horária tem a sua finalidade, ou seja, durante o dia destina-se, principalmente, às pessoas inactivas e a partir das 18H, a todos os elementos da família, apresentando programas que pretendem atenuar a fadiga de um dia de trabalho. Entre as 20H e as 21H, as emissões destinam-se aos pais e procuram satisfazer a globalidade do público. O fim da noite poderá apresentar programas mais instrutivos e culturais. Portanto, cada emissão procura adequar-se à diversidade da população que poderá assistir ao *media* nesse período.

A autora salienta que se verifica a estandardização dos formatos que varia segundo a época e o género e a diminuição de determinados géneros televisivos, nomeadamente do documentário. Apesar do canal RTP1 e as estações generalistas privadas coincidirem ao nível da tipologia de grelha de programação, as motivações intrínsecas aos correspondentes segmentos de audiência-alvo diferem, assim como as linhas de acção estratégica inerente ao conteúdo programático.

B.1.2. As estações televisivas e a contra-programação

De acordo com *Fernandes (2001)*, no Canal 2 a programação denomina-se de compósita, ou seja, mantém a calendarização, semana a semana. Salienta que o seu objectivo não consiste na procura de públicos, mas, sim, na coerência dos programas transmitidos e no facto de existirem algumas particularidades de visionamento segundo as horas. A programação compósita preocupa-se com a audiência, mas, também, com a imagem da estação televisiva e cultiva a totalidade dos géneros. A sua grande diferença face à programação federativa reside no facto da segunda, seguir a procura do público e na compósita ser o responsável de programação quem decide o que se irá mostrar ao público.

A grelha de programas pode funcionar como um meio de concorrência entre as estações televisivas, e que poderá ser classificada em concorrência por mimetismo ou por complementaridade.

➤ Na concorrência por mimetismo, a grelha é construída de maneira aos telespectadores a conseguirem identificar com a estação de televisão, mas, também, em função dos conteúdos transmitidos pelas restantes estações. Num contexto de concorrência em que se procura conquistar o máximo de audiência e em que domina a ideologia da programação federativa, as diversas televisões, opõem os tipos de emissão em horas idênticas, visando o mesmo perfil de público. Este cenário reflecte uma tendência à uniformização da paisagem audiovisual e a uma estandardização da oferta com a transmissão dos mesmos géneros, às mesmas horas e para os mesmos públicos, como, por exemplo, se verifica no decorrer da tarde e a partir do horário nobre.

➤ A concorrência por complementaridade ou contra-programação, caracteriza-se por existir uma preocupação pelos diferentes interesses do público, no sentido de apresentar às mesmas horas que as televisões concorrentes, uma oferta de programas diferentes e a mesma tipologia de programa em horários diferentes. Associa-se ao conceito de contra-programação, o denominado “efeito de herança” que ocorre entre os programas colocados sequencialmente. Assim, por exemplo, a seguir ao espaço informativo da noite as pessoas assistem à telenovela.

De acordo com *Fernandes (2001)*, nos três primeiros anos de televisão privada, em Portugal, a tendência ao nível das estratégias de programação incidia na concorrência por mimetismo. Posteriormente, a contra-programação parecia ser o instrumento eleito para disputar as audiências, sobretudo entre os canais RTP1 e SIC. No horário nobre, por vezes, a RTP1 transmite um jogo de futebol que apresenta uma ampla audiência masculina ou reportagens e concursos e em contra-programação a SIC, defende-se com uma telenovela que apresenta uma audiência feminina fiel a este tipo de género televisivo.

Apesar de ambos os canais privados generalistas, nos períodos de maiores audiências, optarem pelos programas de ficção, a TVI deve o seu rápido crescimento à transmissão de géneros programáticos complementares, como, por exemplo, os *reality shows* e as telenovelas de origem nacional, que alguns anos mais tarde, também, foram adoptadas pela SIC. A exclusividade da SIC para com a transmissão das telenovelas da Rede Globo acabou por penalizar os seus *shares* de audiência, assumindo uma maior relevância as séries de origem nacional. Consequentemente, posteriormente, a concorrência entre a SIC e a TVI acabou por se consolidar numa concorrência por mimetismo, transmitindo-se no horário da noite em ambos os canais algumas telenovelas de origem nacional e, optando o canal SIC por transmitir as telenovelas brasileiras da Rede Globo no final da noite e no pré-prime-time ou após o espaço informativo da tarde. A concorrência entre os canais contribuiu para o desenvolvimento de algumas estratégias de programação, alicerçadas nas audiências dos programas. Actualmente existe uma grande semelhança entre as grelhas dos canais privados, com a transmissão de *talk-shows*, séries, telenovelas, concursos musicais e filmes e apenas na RTP1 se recorre por vezes à contra-programação.

B.1.3. A evolução dos shares de audiências dos canais generalistas

Shares de Audiência				
Anos	RTP1	RTP2	SIC	TVI
1992	72,2	17,9	8,5	
1993	61,5	17,6	14,3	6,6
1994	46,9	9,8	28,4	14,7
1995	38,6	6,4	41,4	13,8
1996	32,6	6,5	48,6	12,3
1997	33,0	5,6	49,3	12,1
1998	31,5	6,2	49,2	13,1
1999	28,5	6,0	48,1	17,4
2000	24,3	5,6	42,2	20,8
2001	20,1	5,6	34,0	31,9
2002	21,1	5,3	31,5	31,4
2003	23,8	5,0	30,3	28,5
2004	24,7	4,4	29,3	28,9
2005	23,6	5,0	27,2	30,0
2006	24,5	5,4	26,2	30,0
2007	29,8	6,2	29,7	34,3
2008	28,1	6,6	29,3	36,0
2009	29,4	7,1	28,5	35,0
2010	24,2	5,3	23,4	27,5

No *Quadro B.1.3.1* consta a evolução dos *shares* de audiência dos quatro principais canais generalistas. Antes da existência dos canais privados, o canal RTP1 não apresentava uma estratégia orientada para o mercado. Consequentemente, rapidamente a SIC conseguiu atingir a liderança. A TVI apresentou-se como a terceira estação televisiva, mas, perdendo ao nível de *shares*, até 1997.

A estratégia inicial da TVI baseava-se num público objectivo, essencialmente, composto por donas de casa e jovens que assistiam à TV no decorrer da tarde. Em 1998, a TVI foi adquirida pela Media Capital e definiu-se uma nova estratégia ao nível da programação e de gestão organizacional, iniciando uma tendência ascendente no seu desempenho global. O horizonte temporal considerado, na presente tese, abrange desde 2000 a 2010 e algumas das análises realizadas incidem nos anos de 2000 a 2004, uma vez que este período se caracteriza como sendo de ruptura face aos anos anteriores com o progressivo acréscimo do share do canal TVI, apresentando valores equiparáveis ao do canal SIC no final deste período.

Quadro B.1.3.1 – Evolução de *Shares* (1992-2010)
Fonte: Marktest (2009)

Em 2000, a TVI ainda apresentava um *share* sensivelmente de metade face ao valor da SIC, 20,8% e 42,2% respectivamente. Este período corresponde a um horizonte temporal que se caracteriza de rotura e de transição face aos anos anteriores e irá ser privilegiado em determinadas análises a realizar para avaliar em que medida a alteração no desempenho dos canais influência a evolução dos indicadores de audimetria e a viabilidade da sua previsão.

O canal TVI cresceu retirando audiência ao canal SIC e à RTP1. Este crescimento deve-se à transmissão de programas adquiridos à Endemol e em especial ao programa “Big Brother”. Para além dos *reality shows*, a aposta da TVI na ficção nacional, também, contribuiu para a maximização das suas audiências, como por exemplo, os programas “Todo o Tempo do Mundo”, “Jardins Proibidos” e “Olhos de água” que iniciaram este ciclo. Desde 2005 a 2007 verifica-se um acréscimo no valor do *share* dos principais canais generalistas, mas mantendo-se a sua ordenação. Desde 1999, a TVI tem apostado nesta tipologia de programas, tendo cada telenovela maiores índices de audiências face às suas predecessoras. Numa reportagem exibida em Setembro de 2010, na SIC, sobre a temática da televisão privada em Portugal, refere-se que o desenvolvimento da ficção nacional resulta dos elevados níveis de analfabetismo que ainda existem actualmente. Logo, este factor inviabiliza uma correcta caracterização da exposição televisiva se não recorrer às tecnologias de informação.

As telenovelas nacionais apresentam-se como um pilar fundamental de crescimento, assim como os *reality shows*. O ano de 2005 apresenta um novo ponto de viragem com a TVI a liderar as audiências (30% face aos 27,2% da SIC). Outro fenómeno de audiências corresponde ao programa “Morangos com Açúcar” sendo este dirigido a um segundo nicho de mercado ainda não alcançado: os adolescentes. Em 2006 assinou um protocolo com a Sport TV, segundo o qual a TVI adquiriu os direitos de transmissão dos jogos de futebol da primeira liga desde 2006 até 2008 inclusive. Os segundos e quartos trimestre correspondem aos que maximizam, os investimentos comerciais. Logo, a TVI relativamente a segmentos alvo realiza o *targeting* de donas de casa desde os 25 aos 45 anos pelo seu valor comercial. Assegura que diferentes programas sejam transmitidos entre a semana e os fins-de-semana, mantendo como prioridade a ficção nacional.

A TVI lidera o horário nobre desde 2001, atingindo um valor médio de 27,5% em 2010. O canal SIC apresenta uma programação claramente especializada em programas mais populares. A sua ligação à Globo influencia a natureza da sua programação, sendo a sua lei a do mercado. A estruturação da grelha de programas visa maximizar o seu *share* de audiência, transmitindo algumas telenovelas brasileiras no decorrer do dia e as nacionais em maior grau no período da noite. Nos últimos anos, também, tem apostado em projectos de origem nacional, em virtude dos bons resultados e da concorrência exercida por parte da TVI.

A TVI de início adoptou uma grelha de programas tão diversificada quanto possível, mas apostando fortemente na programação nacional a qual tem grandes custos associados. No entanto, a pouco e pouco foi alterando a sua estratégia tendo em conta a sua sobrevivência como estação de televisão generalista, verificando-se um acréscimo também na emissão de séries e filmes norte americanos. O interesse crescente do público por séries nacionais é uma mudança muito significativa no respeitante aos hábitos da audiência, que se tem verificado não só em Portugal, como também nos mercados internacionais.

Em 2001, apostou no reality show “Big Brother”, o qual revolucionou o mercado televisivo, sendo o primeiro de sucessivos programas de idênticas características, como por exemplo, “A Quinta”, “O Circo”, “O Meu Odioso Noivo”, entre outros. Na RTP1, o contrato de concessão de serviço público e a necessidade de obter receitas, através da publicidade determinam a elaboração da grelha. A RTP1 tem que gerir estes dois eixos, sendo uma tarefa difícil, pois o público prefere os programas mais populares face aos educativos e culturais. Este canal transmite frequentemente eventos desportivos, galas, o festival da canção, prémios desportivos, assim como, filmes, concursos mais didácticos ou populares e programas de debate e reportagem. Sem dúvida que o seu ponto forte são os programas destinados à informação.

B.1.4. Os canais por TV Cabo em Portugal

Em 1994, a TV por Cabo chega a Portugal, oferecendo dezenas de canais estrangeiros e suscitando o aparecimento de novos canais portugueses por cabo.

Actualmente, existem mais de 10 canais nacionais. A televisão por cabo é, actualmente um meio de acesso à TV massificado, que se encontra em mais de quatro milhões de lares e que equivale a 1,5 milhões de assinantes.

A evolução mais acentuada verificou-se entre 1995 e 2002, sendo mais gradual após este período (*Quadro B.1.4.1*).

Anos	Alojamentos	Assinantes
1995	377	58
1996	977	171
1997	1466	383
1998	1827	596
1999	2259	760
2000	2601	935
2001	3024	1119
2002	3349	1261
2003	3494	1335
2004	3631	1343
2005	3773	1400
2006	3825	1421
2007	4040	1489
2008	4275	1475
2009	3990	1452
2010	4033	1437

Quadro B.1.4.1 – Evolução do número de alojamentos E do número de assinantes de TV por Cabo (1995-2010). Valores em Milhares
Fonte: Anacom (2010)

Embora o serviço de distribuição de televisão por cabo tenha arrancado em 1994, no que se refere ao processo da medição das audiências este apenas foi implementado em 1999, com uma amostra de 250 lares, representativos do universo Cabo.

Em 2003, de modo a acompanhar o crescimento da taxa de penetração deste serviço, a Marktest Audimetria alterou o procedimento da medição das audiências dos lares com cabo, abandonando uma amostra fixa, e passando, a partir de Setembro, a ter uma amostra proporcional ao universo cabo. Os dados de audimetria disponibilizados referem-se a uma amostra com 250 lares com TV Cabo. A evolução da quota da audiência do conjunto dos canais via cabo tem manifestado uma progressão positiva, registando, na média do primeiro semestre do ano 2004, um share de 11,4%. Em 1999, esse valor não ia além dos 2,9%.

Apesar da tendência crescente dos *shares* de audiências dos canais por TV Cabo, o interesse das agências de planificação de meios no decorrer do período de 2000 a 2007 incide essencialmente nos canais generalistas, e em especial na RTP1, SIC e TVI. Nas bases de dados de audimetria constam os índices de audiências de canais por Cabo, mas não consta a descrição da programação. Com base nas análises de duplicação a realizar podemos confirmar se a exposição a estes canais realmente penaliza significativamente a evolução dos principais indicadores de audimetria que quantificam o desempenho dos canais generalistas.

O telespectador de canais Cabo caracteriza-se por pertencer aos extractos sócio-económicos mais elevados, classe alta/média alta e classe média e, logo, a alguns dos segmentos que apresentam uma menor exposição ao *media*. Por exemplo, a classe D apresenta apenas 13,8% para o ano de 2004 e a Classe AB de 29,6%. No entanto, é de salientar que a quota destes segmentos tem vindo, ligeiramente, a perder importância. No que se refere ao género, o telespectador dos canais Cabo é fundamentalmente masculino, embora a importância da população feminina tenha crescido, facto a que a consolidação da oferta de conteúdos dirigidos a esta população, como é o caso da SIC Mulher, não serão alheios.

Em termos evolutivos, é na caracterização por grupos etários que se observam as maiores flutuações, ou seja, enquanto que nos três primeiros anos em que passámos a dispôr de dados de audiências, o telespectador Cabo era maioritariamente constituído por indivíduos entre os 4/34 anos, assiste-se a um significativo incremento da faixa 35/54 anos em detrimento das faixas mais jovens.

Os conteúdos generalistas nos canais por cabo lideram as quotas de audiências, facto a que o nível de oferta não será de todo alheio. Por outro lado, a importância que a tipologia de Informação ganhou a partir de 2001, deve-se essencialmente ao aparecimento da SIC Notícias, pois relativamente à oferta, desde 1999 até 2004, foram sempre contabilizados 7 canais.

B.1.5. Perspectivas futuras para o sector televisivo

A tipologia Infantil onde se enquadra o canal codificado Disney Channel tem mantido a 3ª ou 2ª posições no ranking de share por tipo de conteúdo, embora o valor de quota tenha sofrido algumas oscilações ao longo dos anos. Por sua vez, os canais cujos conteúdos temáticos são Documentários/Cultura têm vindo a perder expressão, desde 2000, tal como acontece com os canais de Cinema, sendo de referir que na actual oferta cabo, três dos quatro canais desta tipologia são Pay TV. Em termos da tipologia Desporto, observa-se uma nítida estabilidade após 2001. Nos conteúdos de Música regista-se uma quebra gradual nos valores de share até 2002, contudo, a partir de 2003, assiste-se a uma retoma, facto a que as alterações ocorridas no canal MTV, que passou a emitir em Português, não serão de ignorar, visto que a oferta de canais desta tipologia não se alterou.

As previsões de longo prazo apresentam uma maior complexidade, sendo difícil conseguir prever como a audiência poderá reagir à introdução de um novo programa ou tipologia programática e a alterações de gestão estratégica de cada uma das estações televisivas e dos canais. Contudo, as análises realizadas aos indicadores acumulados de audimetria permitiram concluir que existe uma maior influência estrutural na exposição televisiva, o que poderá aumentar a previsibilidade dos comportamentos futuros.

Segundo *Faustino (2007)*, a indústria televisiva encontra-se influenciada, essencialmente, por seis factores, nomeadamente: a perda de *share*; as agências de publicidade, a concentração existente no mercado, uma elevada estrutura de custos associada à grelha de programação, a TV Cabo e a “TV Reality”.

Os indicadores de audimetria, sem dúvida, que condicionam a estruturação das grelhas, assim como o número de canais, sejam eles generalistas, ou não, e os agentes que financiam a publicidade, os custos relacionados com a aquisição e produção dos programas e o forte interesse que os “reality shows” têm despertado no público.

A estratégia delineada pela TVI tem sido frutífera, desde 2000, permitindo-lhe atingir a liderança desde 2005 e deverá manter a sua posição desde que continue a apostar na ficção nacional, nos “reality shows” e nos eventos desportivos. Sendo o mercado televisivo muito volátil e vulnerável a alterações, qualquer estratégia de cariz inovador ou direccionada a um nicho de mercado, ainda não identificado, poderá reverter esta situação e alterar o desempenho de cada um dos canais, assim como a privatização de um canal da RTP . “... ainda assim a empresa (TVI) não é imune às alterações verificadas no ambiente externo e em 2009, ano de crise económica, o decréscimo da publicidade levou a uma redução dos lucros da estação.” (Ferreira, Reis & Santos, 2011, p.11).

De acordo com *Aris e Bughin (2005)*, a TVI não apresentava na altura uma grande margem de crescimento e os canais não generalistas ou a RTP1 poderiam vir a exercer alguma concorrência e penalizar o desempenho dos canais privados generalistas. Salientam que o futuro da SIC apresentava-se como incerto:

Growth opportunities in TVI’s current free-to-air business are limited. Its’ market share is under threat due to increasing audience fragmentation, partly due to new audience platforms. Portuguese cable penetration is growing, especially in affluent cities and in long run, digital terrestrial television may also be a threat ... In addition, RTP is increasingly becoming a real competitor, after the changes occurred in the public broadcasting concept. The future role of SIC is also still unknown. Therefore TVI will have to address certain issues proactively in the future. (p.244)

Verificámos que a TVI, desde 2005, apresentou uma tendência ascendente e, posteriormente, em sentido inverso, assim como os restantes canais. Em *Faustino (2007)*, podemos encontrar uma análise SWOT realizada à empresa Media Capital. Desde 2005 que a TVI mantém a liderança e a RTP1, também, mantém a sua posição desde 2003. Apenas a SIC tem apresentado um maior decréscimo nos seus *shares* de audiência no decorrer deste período. Determinadas previsões são difíceis de realizar mesmo quando se recorre a métodos qualitativos e deve-se ter em atenção o carácter dinâmico deste sector. As previsões devem ser realizadas a curto e médio prazo e os resultados reavaliados quando se verificam consideráveis alterações ao nível da grelha de programação.

A privatização de um dos canais da RTP, também, poderá provocar alterações ao nível do desempenho global de cada uma das estações televisivas. Consequentemente, do ponto de vista da previsão o mais relevante corresponde na realidade à identificação das metodologias mais adequadas e que são função dos contextos, sendo posteriormente necessária uma contínua análise da evolução dos dados e a actualização dos resultados.

B.2. A TELEVISÃO COMO MEIO PUBLICITÁRIO

B.2.1. Os motivos para valorizar o investimento publicitário no sector televisivo

Em relação aos restantes meios de comunicação, a publicidade na TV apresenta múltiplas vantagens, entre as quais, se destacam as seguintes:

- Um elevado alcance sobre a totalidade da população, o que permite uma fácil cobertura de todos os segmentos-alvo. A televisão corresponde ao único *media*, que permite alcançar praticamente a totalidade da população, viabilizando seleccionar de forma eficaz diferentes perfis populacionais e conseguir rapidamente elevadas coberturas sobre donas de casa, adultos, jovens e outros segmentos que apresentam valor comercial. Trata-se de grupos populacionais genéricos sobre os quais cada anunciante poderá realizar ajustes em função dos seus perfis comunicacionais.
- Um reduzido custo por contacto em comparação com os restantes *media*. A televisão permite contactar de uma forma mais económica com o público-objectivo, mas será necessário um elevado investimento para conseguir construir campanhas que ultrapassem o limiar de eficácia mínima. Isto poderá ser explicado com base na própria estrutura e condições de interacção da audiência com o *media* e que torna necessário um número muito superior de contactos face ao que se verifica, por exemplo, em jornais ou revistas.
- Uma elevada qualidade de medição das variáveis publicitárias. A relevância quantitativa deste *media* para o conjunto dos sectores que dele dependem como meio de divulgação dos seus produtos permitiu o desenvolvimento máximo da investigação das audiências.

- O carácter audiovisual do *media* confere-lhe inúmeras vantagens, permitindo aos que elaboram os anúncios explorar ao limite as suas possibilidades convertendo os produtos e as inserções publicitárias num espectáculo em si mesmos.
- A publicidade nos *media* permite que os anunciantes controlem a mensagem publicitária que oferece uma mensagem simples, forte e única. A publicidade funciona segundo o princípio da repetição, ou seja, da multiplicação dos contactos e da continuidade da campanha.
- A publicidade permite reforçar a notoriedade de um marca; difundir informação que pode modificar o comportamento dos segmentos-alvo e associar a uma marca características que incentivem à sua aquisição.

B.2.2. Os factores e possibilidade de negociação na TV: As condições de contratação

A compra do espaço e tempo comercial tem evoluído no decorrer do tempo. As opções ao nível do *media* têm-se fragmentado com o acréscimo no número de estações televisivas e de revistas. Os anunciantes pressionam as agências no sentido de se alcançar uma maior eficácia, uma vez que os custos apresentam uma tendência crescente. O tempo publicitário é vendido em unidades de tempo, sendo frequentes os espaços de 30 segundos. As unidades que não são vendidas constituem o inventário disponível.

As estratégias de preços são função das audiências totais, da tipologia de programas transmitidos e dos segmentos sociodemográficos que caracterizam o seu público. Destes factores, os autores, salientam a existência de cinco efeitos tangíveis no preço negociado e que a estrutura de preços encontra-se, essencialmente, influenciada pelas audiências actuais e históricas, tipologias de programas que oferecem e dos factores de imagem das estações televisivas. Os restantes factores que influenciam o processo de negociação correspondem ao conflito, poder, influencia interpessoal e relações existentes, entre outros de cariz mais subjectivo (McMellom & Eftimov, 2004).

A venda do espaço comercial pode ser realizada em diferentes momentos e apenas na primeira fase designada por “upnt market” existe o compromisso de garantir determinada exposição ao plano contratado. No “scatter market” ou no “spot market” não são dadas garantias por parte das estações televisivas ao nível dos níveis de audiências que podem ser alcançados e quanto à sua composição sociodemográfica. Os preços no “scatter market” podem ser superiores ou inferiores aos da primeira fase de venda do espaço comercial em função das condições do mercado.

Se os *ratings* de determinado canal forem inferiores ao esperado, poderá levar a uma redução do inventário disponível para venda no “scatter market” o que inevitavelmente acarreta uma súbita dos preços. As inserções publicitárias no “up-front market” podem ser vendidas em planos de dúzias e centenas de unidades que cobrem meses, vários dias da semana e tipologias de programas e que são contratadas para cada produto. As estações televisivas apresentam para cada plano um preço e um *rating* sociodemográfico total que deve ser garantido. Por vezes, o preço é dado de forma global, contudo as agências apresentam algum conhecimento sobre o custo associado a cada inserção contratada. A maior parte das inserções, por volta dos 70%, é vendida nesta fase.

Para a construção de planos de *media* é relevante a quantificação da exposição de um colectivo face a um conjunto de suportes e em alternativa será necessário a previsão de valores de *ratings* dos suportes a incluir nos planos. A venda do espaço comercial, também é realizada em função dos valores de GRP mínimos a garantir para o pacote de anúncios que, por sua vez, resulta do somatório dos valores de *ratings* para cada inserção contratada ou do produto da frequência e da cobertura alcançada.

As estações televisivas devem garantir determinado GRP para o plano adquirido pelos anunciantes. Se alguns programas apresentarem valores reduzidos de *ratings* e outros valores superiores face ao esperado, desde que o valor global de GRP supere o valor mínimo acordado, a estação televisiva cumpriu o acordo que foi estabelecido. Caso contrário, deverá compensar o diferencial através da cedência de novas inserções publicitárias.

À medida que as inserções são transmitidas, também são debitados dos anunciantes os custos que lhe estão associados por parte das estações televisivas. Os valores são definidos com base nos valores de *ratings* que estas atingem e em função dos registos internos dos valores de CPM (“demographic cost per thousand”). Se os valores de GRP acordados não forem atingidos o preço não é reduzido e os anunciantes são compensados com unidades adicionais até se alcançar o número de exposições mínimo acordado.

Se o valor for excedido não existe qualquer compensação para o canal e o plano deverá ser mantido. Os canais neste contexto perdem dinheiro por terem subestimado a sua real capacidade de exposição ao segmento-alvo da campanha publicitária, contudo apenas, pontualmente, se costuma verificar esta situação. As unidades adquiridas podem ser reposicionadas algumas vezes antes da sua transmissão, como consequência das alterações horárias de determinados programas, programas especiais ou eventualmente até canceladas se os programas apresentarem valores de *ratings* reduzidos. As alterações realizadas na grelha de programas requerem alterações no posicionamento das inserções e que será realizada de comum acordo entre as estações e os anunciantes ou as agências. Dia a dia, são recolhidas algumas estatísticas a respeito da exposição televisiva dos diversos programas. Os anunciantes podem retirar inserções parcialmente ou a nível global sem serem penalizados e deverá ser definido um preço justo para essas inserções.

Os anunciantes podem simplesmente referir que determinado programa apresenta uma temática que não é aceitável. Uma estação de televisão tem como objectivo não proporcionar aos anunciantes níveis de exposição superiores ao estabelecido. Os preços não estão relacionados com o número de inserções, mas, sim, com o valor do GRP global da campanha. Logo, se os valores de GRP não forem alcançados será suficiente atribuir um maior número de inserções aos anunciantes e as estações *a priori* já esperam ter que lhes ceder unidades adicionais. As unidades devem ser distribuídas entre programas com diferenciadas características.

As agências de meios devem dispor de estimativas de valores de *ratings* e de GRP para garantir os melhores resultados em função do investimento a ser realizado. A partir das previsões realizadas seleccionam-se as combinações óptimas de blocos comerciais que será mais rentável para o seu público-objectivo, sendo necessário posteriormente avaliar os reais resultados alcançados pela campanha. Como passo prévio a qualquer sistema de optimização do investimento é necessário conhecer a audiência dos programas e blocos comerciais onde se possa investir e as previsões devem ser realizadas da forma mais rápida possível.

B.3. A ACTIVIDADE DE PLANEAMENTO DE *MEDIA*

B.3.1. O planeamento de *media* e os “Estudos de audiências”

O que se denomina de “investigação ou estudos de audiências” engloba especificamente o estudo de questões relacionadas com a exposição aos meios de comunicação e apresenta como um dos seus objectos de estudo, as necessidades de informação inerentes ao planeamento de *media*. Verifica-se a necessidade de compreender os comportamentos de consumo televisivo, segmentar as audiências e a quantificação da exposição para o segmento-alvo de uma campanha publicitária. Na actividade de planeamento de *media* existe a necessidade de desenvolver modelos explicativos e/ou de previsão das audiências televisivas que podem ser quantificadas através de inúmeros indicadores.

Os modelos explicativos da interacção do painel com o meio televisivo ou dos índices de audiências podem contribuir para as análises de previsão. Nesta área nunca se poderá definir como objectivo a realização de previsões deterministas, mas, sim, a identificação e análise das dimensões que condicionam maioritariamente a exposição e a evolução dos indicadores que condicionam a venda do espaço comercial. Apenas após a avaliação da influência de determinantes estruturais e individuais na exposição televisiva e na evolução dos principais indicadores de audimetria se poderá definir uma sequência metodológica que permita contribuir para a temática da previsão das audiências televisivas em função dos diversos contextos de análise.

Na análise dos comportamentos de consumo televisivo é relevante compreender quais são os factores que determinam a exposição, uma vez que estas análises condicionam a estruturação da grelha e os conteúdos a incluir nos planos de media. Os modelos explicativos e de previsão de valores de *ratings* programáticos e comerciais ou da exposição televisiva corresponde a uma área crucial para a gestão das empresas ligadas a este sector de actividade.

B.3.2. A evolução dos métodos de medição das audiências

Inicialmente os estudos de audiências eram realizados recorrendo aos inquéritos que proporcionam uma visão estática da exposição televisiva e que proporcionam informação que carece de objectividade. As frequências de exposição neste contexto são calculadas com base na probabilização dos comportamentos e com os erros que, inevitavelmente, acarretam (Zarate, 1997). Os instrumentos tecnológicos de medição de audiências variam substancialmente de país para país.

Na actualidade, os instrumentos que mais se destacam correspondem aos estudos de memória, aos estudos coincidentais, a audimetria e o *audio-matching*. Os audímetros e o *audio-matching* permitem uma análise diacrónica da exposição, produzindo um vasto conjunto de dados que caracterizam a exposição de forma mais detalhada, rápida e precisa. As restantes abordagens são utilizadas essencialmente como instrumentos de validação dos resultados. A principal função destes instrumentos corresponde ao da quantificação das audiências.

Por vezes, também, é necessário recorrer e complementar as análises com dados mais qualitativos de caracterização da exposição televisiva. Os audímetros podem permitir que sejam colocadas questões ao painel, através do desenvolvimento da audimetria interactiva e em alternativa em determinados países recorre-se aos painéis de reacção, onde são colocadas questões a respeito da atractividade exercida por parte de determinados conteúdos. Por vezes esta informação permite a construção de índices que são utilizados em análises de previsão, mas que apresentam consideráveis margens de erro. No *Apêndice A.1* caracteriza-se em maior detalhe as diversas tecnologias de medição das audiências televisivas.

B.3.3. Os pressupostos em que se baseia a investigação das audiências pré-audimetria

O planeamento de *media* realizado com inquéritos pretendiam verificar se se produz o contacto entre um suporte e um indivíduo e qual a sua frequência. Corresponde a uma informação essencial para a construção dos planos e que se baseia na utilização da informação histórica para prever os comportamentos de curto prazo durante o horizonte temporal que dure uma campanha. Será necessário conhecer a audiência dos programas em que se vão inserir os anúncios. O pressuposto de base corresponde a aceitar a regularidade dos comportamentos, ou seja, a existência de ciclicidade, de repetição, de semana para semana e/ou de dia para dia. Neste contexto aceita-se que os indivíduos não modificam os comportamentos de forma significativa se não existirem consideráveis alterações na grelha de programas e o número de pessoas que contactam com determinado suporte considera-se constante, assim como a sua composição sociodemográfica.

B.3.4. O contributo da audimetria para o planeamento de *media*

Recorrendo à audimetria podemos realizar com rapidez uma análise diacrónica da audiência para uma elevada dimensão de suportes e considerando uma amostra que se considera representativa do universo televisivo. Permite avaliar de forma objectiva como o painel interage com o televisor e analisar comportamentos de consumo televisivo que decorrem da evolução verificada na oferta televisiva e da tecnologia, como por exemplo, o “zapping” e o “zipping”. Quantifica-se a exposição ao segundo o que permite o cálculo automático de uma grande diversidade de indicadores, entre os quais, as audiências acumuladas e a duplicação de audiências que não seria viável conhecer recorrendo aos inquéritos que apenas proporcionam uma visão estática das audiências e subjectiva. No *Apêndice A* caracteriza-se em maior detalhe o processo de audimetria.

B.3.5. A elaboração de uma campanha publicitária

B.3.5.1. Os actores e as suas funções

Os principais actores intervenientes neste sector correspondem às estações televisivas, às agências de planificação de meios, às agências de publicidade e aos anunciantes. Entre os anunciantes e os *media* existem vários intermediários cujas principais funções se organizam em torno de três áreas: a concepção das campanhas, a realização das campanhas e a venda do espaço comercial. As agências de publicidade, inicialmente, correspondiam a empresas correctoras de espaço, mas no decorrer do tempo foram progressivamente proporcionando outro tipo de serviços aos seus clientes, nomeadamente realizar um conselho estratégico, conceber mensagens e planos de *media* e por último comprar espaço publicitário e produzir as campanhas.

As centrais de compra correspondem a organizações que numa primeira fase apenas promoviam a compra de volumes de espaço muito significativos, conseguindo preços mais vantajosos para os anunciantes. Posteriormente, também desenvolveram uma perspectiva de serviço, e entraram no campo de planeamento de *media*, surgindo as empresas designadas por agências de meios. As agências de meios definem a selecção de suportes, formatos e número de inserções e numa fase posterior realizam o acompanhamento e a avaliação dos resultados da campanha. Para além do planeamento, a compreensão do alvo e do respectivo perfil de exposição aos *media* também é determinante.

B.3.5.2. As etapas e as responsabilidades de uma campanha

A elaboração de uma campanha publicitária inclui quatro etapas fundamentais e principais, subdivididas em várias tarefas, nas quais os dois principais actores, o anunciante e a agência intervêm com pesos variáveis: a redacção do briefing; a concepção da campanha; a execução da campanha e a avaliação *a posteriori* da sua eficácia.

Uma agência de meios corresponde a uma empresa dedicada profissionalmente à investigação dos *media* e suportes, à planificação da estratégia de espaços e tempos, compra e verificação da comunicação. Têm como função realizar o planeamento de *media* que consiste em escolher os *media* e suportes, ou seja, os veículos da publicidade, nos quais devem estar inseridos os anúncios publicitários. Seleccionam o conjunto de meios, a sazonalidade e a pressão da campanha que se deve inspirar em três objectivos gerais: a cobertura do alvo, a repetição e a qualidade da comunicação.

Os suportes seleccionados devem permitir alcançar o maior número possível de pessoas pertencentes ao alvo seleccionado e com um valor adequado de frequência. Nesse sentido a duplicação de audiências entre suportes corresponde a uma das necessidades de informação essencial para a construção das campanhas publicitárias.

A implementação de um plano de *media* comporta quatro etapas sucessivas. Inicialmente desenvolvem a estratégia de *media*, ou seja, a selecção do *media*, formato e período do ano. Numa segunda etapa a preparação do plano de *media*, que consiste em escolher para cada *media* os suportes específicos e fixar o calendário exacto das inserções. Segue-se a compra do espaço e o controlo, verificando se as inserções estão em conformidade com o plano adquirido. Para formular um plano de *media* será necessário dispôr de um grande número de informações de base sobre a audiência dos principais suportes, como por exemplo, indicadores que quantificam a exposição das audiências, assim como informação mais qualitativa.

B.3.5.3. A estratégia de *media*

Uma vez estabelecidos os objectivos e determinado o público-alvo, a seguinte etapa corresponde à formulação de uma estratégia de *media*, que visa identificar a melhor combinação de meios para que a mensagem publicitária chegue eficazmente ao alvo seleccionado. A estratégia pode ser definida como o conjunto de métodos utilizados para alcançar este objectivo, tendo em conta o entorno.

O planificador deve realizar uma análise dos *media* relativamente ao seu perfil de audiências e a sua afinidade para com o *target*, valorizando aspectos quantitativos e qualitativos que aconselham ou desaconselham a sua utilização na campanha, assim como o papel que deve desempenhar cada um dos *media* seleccionados no conjunto do plano delineado.

B.3.5.4. O plano de *media*: A táctica

Um suporte designa um programa ou faixa horária na qual se irá inserir o anúncio. Um plano de *media* corresponde à selecção dos suportes que possa garantir a maior cobertura do mercado alvo e a determinada frequência.

Verifica-se a necessidade de avaliação da variabilidade dos principais indicadores quando associados a programas ou faixas horárias e, assim, definir quais são os suportes que devem ser considerados no planeamento de *media*. As faixas horárias podem apresentar uma maior estabilidade. Logo, existe a necessidade de segmentar a emissão diária em função da evolução das variáveis critério. Alterações de horários que não pertençam ao mesmo agrupamento temporal podem condicionar em maior grau a posterior quantificação da exposição. Verifica-se a necessidade de segmentar uma semana e dia em unidades menores, mais operativas para o anunciante. Posteriormente será necessário estimar quantos entram em contacto e o seu perfil, entre outras variáveis.

A audiência de um suporte corresponde ao número de pessoas que entram em contacto durante determinado horizonte temporal. Contudo, existe uma grande diversidade de indicadores que a caracterizam.

A audiência define-se como o número de indivíduos que declaram entrar em contacto com um media ou suporte, num determinado período de tempo. Em função do estudo de audiências, podem-se considerar diferentes indicadores, uma vez que cada estudo pretende avaliar um período de tempo e utiliza um método, universo e amostra diferente.

Os indicadores de audiências correspondem a uma moeda de troca para os diversos suportes, condicionando as suas tarifas. Logo, o seu conhecimento e a sua previsão são indispensáveis quando se têm que tomar decisões referentes à aquisição de espaço comercial. Atendendo à elevada dimensão do número de suportes que existem em determinada semana e considerando a totalidade dos canais, não é viável a quantificação das audiências para a totalidade. Consequentemente, será necessário definir critérios de selecção que se baseiam nas frequências acumuladas de contacto e na duplicação das audiências.

Será necessário seleccionar os suportes que permitam uma maior repetição da exposição ao público alvo e calcular os valores de *ratings* médios ou em alternativa a cobertura do alvo e o número médio de contactos que possa garantir determinados valores de GRP para a campanha publicitária.

A selecção de um plano de *media* apresenta grande relevância do ponto de vista financeira, uma vez que os custos da compra de espaço são muito elevados. O plano de *media* consiste em escolher os suportes nos quais vão estar inseridos os anúncios publicitários que possa garantir a cobertura que se estabeleceu do mercado alvo. Esta selecção é realizada considerando alguns critérios quantitativos, entre os quais, a cobertura do alvo, a penetração e composição da audiência, a afinidade da audiência, o tempo médio de contacto do indivíduo face ao suporte, as taxas de repetição, o nível de saturação publicitária, os custos das inserções, a flexibilidade da contratação e características técnicas do media: cor, espaço, movimento.

Os suportes são seleccionados de forma a permitir expor a mensagem ao segmento visado com uma frequência suficiente para as fazer agir, ou seja, para favorecer a aquisição dos produtos e serviços divulgados. Consequentemente, as taxas de repetição, ou seja, a duplicação de audiências apresenta grande relevância. Devem, também, permitir atingir o maior número de pessoas pertencentes ao alvo definido, existindo uma grande diversidade de indicadores, sendo um dos mais relevantes a variável designada por *ratings*.

Etapas da Estratégia de Meios

- (1) Valorização dos medias em função de factores quantitativos e qualitativos
- (2) Suportes Seleccionados, justificação da sua selecção e papel desempenhado
- (3) Distribuição de tempos
- (4) Percentagem de Orçamento destinado a cada suporte por período

Numa primeira fase seleccionam-se os *media*, formatos e períodos do ano. Neste contexto tem-se em consideração critérios quantitativos e qualitativos, como por exemplo, a cobertura do alvo, o poder de repetição, os custos, prazos e a acessibilidade. Segue-se a elaboração do plano que consiste em escolher os suportes específicos em cada media e fixar o calendário exacto das inserções.

Quadro B.3.5.4.4.1 – Etapas da Estratégias de Meios

Fonte: Elaborado pela autora

Os suportes televisivos são seleccionados em função da periodicidade, alcance, afinidade, rentabilidade, imagem, credibilidade, posicionamento, entre outras características.

Quando se seleccionam os suportes é necessário distribuir o orçamento total tendo em conta o papel atribuído a cada um deles e os custos associados. Os suportes são hierarquizados atendendo aos critérios definidos e são seleccionados atendendo a todos os factores que podem condicionar a planificação. Uma vez seleccionados os suportes em função dos factores referidos anteriormente e de outros factores estabelece-se a distribuição das inserções, existindo vários tipos de formas de distribuição, como por exemplo, a publicidade contínua, a intermitente contínua, a intermitente crescente ou decrescente, a rítmica, a estacional, a crescente e a decrescente. Finaliza-se com o calendário das inserções, o resumo do valor a afectar ao plano e a sua avaliação.

Na fase de avaliação do plano recorre-se a uma série de indicadores, entre os quais a cobertura do alvo, o OTS, o GRP, os contactos conseguidos, o valor do investimento e o custo por GRP da campanha. Alguns destes indicadores caracterizam-se no seguinte capítulo. Uma vez definido o plano final segue-se a compra do espaço e por último a fase de controlo durante a qual se verifica se as inserções estão em conformidade com o que foi adquirido.

Para se seleccionar os suportes a incluir num plano de *media* é necessário registrar cada um dos suportes previsíveis em três escalões:

- O Escalão Força corresponde à ordem pela qual se classificam os suportes do ponto de vista da audiência útil, ou seja, quanto ao número de pessoas pertencentes ao público-alvo que são expostas a uma inserção nesse suporte. Este indicador diferencia-se da audiência instantânea que corresponde ao número de pessoas que são expostas a uma inserção no suporte considerado;
- O Escalão Afinidade mede a correspondência entre a audiência de cada suporte e o alvo publicitário e corresponde a um dos indicadores sociodemográficos que são disponibilizados nas bases de dados de audimetria;
- O Escalão do Custo por contacto útil permite a classificação dos suportes em função do seu custo por contacto.

B.3.5.5. Os modelos de *media*: A exposição, a optimização e a avaliação dos efeitos da publicidade

Na actividade de planeamento de *media* são necessários três tipologias de modelos; os modelos de exposição, os modelos de optimização e de avaliação dos efeitos da publicidade. Os **modelos de exposição** têm como função quantificar a exposição televisiva, sendo essencial determinar se os telespectadores contactam com os suportes e com que frequência para se poder calcular as audiências acumuladas, ou seja, o número de telespectadores expostos aos suportes após sucessivas transmissões dos mesmos. Após a quantificação e previsão das audiências os planos são construídos através da aplicação de **algoritmos de optimização** que pretendem maximizar a exposição ao plano delineado, mas considerando determinadas restrições que podem ser económicas ou relacionadas com determinadas características dos suportes e que devem ser respeitadas quando são construídos os planos. Por último os **modelos dos efeitos da publicidade** pretendem avaliar em que medida a exposição às campanhas publicitárias incentiva a aquisição dos produtos e serviços divulgados.

B.3.5.6. A necessidade de desenvolver análises explicativas e de previsão das audiências

A investigação das audiências representa um conjunto de metodologias que tem como finalidade o conhecimento das características do receptor, que forma parte do processo de comunicação, sendo esta uma informação indispensável para que os planificadores possam tomar decisões fiáveis, contribuindo, assim, para uma maior eficácia das estratégias de comunicação publicitárias.

Nos anos 60, o mercado publicitário enfrentou grandes dificuldades derivadas da ausência de informação, o que levou os anunciantes, os meios de comunicação e as agências a tomar iniciativas que levaram ao desenvolvimento de novos instrumentos de medida que viabilizam a extracção de informação essencial à sua gestão, como por exemplo, a tecnologia de audimetria. Actualmente, persistem inúmeras questões em aberto no âmbito da análise da exposição televisiva e da construção dos planos de *media*, em virtude do desenvolvimento da Internet, da proliferação dos suportes exteriores, do desenvolvimento da tecnologia digital, entre muitos outros factores.

Os estudos de *media* e de audiências de TV constituem o *input* de todo o trabalho desenvolvido, quer para a definição da estratégia de comunicação, quer para o planeamento das campanhas. Os estudos de audiências lidam com questões relacionadas com a exposição televisiva, a estimação e o conhecimento de onde, como e quando investir para rentabilizar o investimento. Através da análise das audiências televisivas torna-se possível a selecção dos suportes mais adequados aos objectivos do anunciante e através das variáveis de audiências realiza-se o acompanhamento das campanhas.

A forma como os telespectadores interagem maioritariamente com o meio televisivo e os programas condiciona as abordagens de previsão da exposição televisiva que são mais adequadas. Existe uma forte diversidade de contextos explicativos e de previsão que dependem, por exemplo, do público objectivo, faixas horárias, indicadores em análise e nível de agregação dos dados, o canal, mês, entre outras variáveis.

A possibilidade de realizar uma indução depende da existência de regularidade na série de acontecimentos e comportamentos observados. Logo, a caracterização da exposição televisiva sob algumas vertentes e a compreensão dos factores que a influenciam será essencial para conseguir promover melhores resultados no âmbito da previsão das audiências televisivas e do planeamento de *media*.

Do ponto de vista da indústria televisiva é determinante segmentar as audiências e caracterizar, explicar e prever os índices de audiências programáticos e comerciais ou a exposição televisiva. Esta informação permite às estações televisivas a fixação das tarifas publicitárias, a avaliação do desempenho dos programas e canais, assim como a estruturação da grelha de programas. Em relação às agências de meios permite-lhes a maximização do alcance das campanhas em função do seu orçamento, o targeting de determinados segmentos sociodemográficos e a compra de espaço comercial.

Podemos definir a planificação de *media* como a optimização de uma combinação de suportes e *media* com um custo económico que permitia atingir os objectivos de publicidade previamente definidos. Destaca-se o difícil equilíbrio entre os objectivos delineados e o seu custo económico. A complexidade de combinar os *media* e os suportes, reside em que, apesar da importante componente Matemática utilizada, o conhecimento do mercado será o que permite, a partir de orçamentos e objectivos similares, o desenvolvimento de táticas de media com escassos pontos em comum.

A televisão corresponde ao *media* privilegiado pelos anunciantes, conseqüentemente um plano em geral desenvolve-se a partir de uma estrutura central baseada na combinação de suportes televisivos e, posteriormente, podem ser realizados alguns ajustes com a incorporação de outros meios secundários.

APÊNDICE C

A INVESTIGAÇÃO REALIZADA NO SECTOR TELEVISIVO E ENQUADRAMENTO DA TEMÁTICA PROPOSTA

O sector televisivo apresenta uma grande diversidade de âmbitos de análise, que recorrem a diferentes disciplinas, instrumentos, fontes de dados e métodos de investigação. Neste capítulo pretende-se realizar uma síntese das principais áreas de investigação, permitindo, assim, o enquadramento das análises propostas na presente tese.

C.1. AS ORIENTAÇÕES TEÓRICAS, OBJECTOS, DISCIPLINAS E ÁREAS DE ANÁLISE

A televisão é o meio audiovisual que tem suscitado um maior interesse no decorrer das últimas décadas, mantendo, actualmente, a sua liderança. Os objectivos, disciplinas, instrumentos, orientações e motivações que podem ser identificados na investigação realizada neste sector são inúmeros. Na actualidade, proliferam os estudos e as investigações multidisciplinares, existindo trabalhos psicológicos, sociológicos, financeiros, empresariais, entre muitos outros, que oferecem a sua particular visão deste sector com a finalidade de aprofundar o fenómeno dos meios de comunicação e das suas audiências (*Casetti & Chio, 1997*). As publicações existentes podem ser diferenciadas em análises qualitativas versus quantitativas (*Stavitsky, 2000*); empresariais (*Sueiras, Morales & Ibanez, 2007*) ou académicos (*Cooper & Tang, 2009; Gensh & Sharman, 1980*), assim como associados a diversas áreas científicas que abordam questões diferenciadas, recorrendo, consequentemente, também, a diferentes metodologias de investigação.

Para se conseguir delinear um quadro global das diferentes formas de estudar este *media* será necessário considerar as inúmeras disciplinas que concorrem para a sua análise. Os objectivos das análises determinam quais devem ser as disciplinas implicadas, instrumentos adoptados, a perspectiva segundo a qual observam as questões e a direcção que se imprime à investigação (Cassetti e Chio (1997). A justaposição de diferentes perspectivas pode conduzir a um problema de integração.

Segundo Cassetti e Chio (1997), os núcleos temáticos abrangem: a análise da produção e a análise da oferta televisiva, isto é, dos programas em concreto e das grelhas de programação, e por último o consumo televisivo. Os autores referem que quando o objecto de estudo corresponde ao telespectador podemos identificar inúmeras vertentes de análise, entre as quais as seguintes: (1) a análise dos índices de audiências, do número e identidade dos telespectadores; (2) a análise das selecções realizadas, ou seja, das motivações, orientações e preferências que condicionam a exposição a programas e espaços comerciais; (3) a análise dos processos de compreensão e de valorização; (4) a análise dos efeitos a nível cognitivo e comportamental a curto e longo prazo e por último (5) a análise da dieta de *media*, do modo como o consumo televisivo se relaciona com o consumo cultural. As análises a desenvolver enquadram-se no primeiro contexto de análise, que abrange o desenvolvimento de modelos explicativos e de previsão das audiências, sob diversas vertentes, recorrendo à investigação quantitativa e a fontes de dados designadas por “*ratings*” (Stavistky, 2000).

O contexto de análise das selecções realizadas não abrange apenas a análise das motivações e das preferências, mas, também, por exemplo, a influência das variáveis estruturais e dos comportamentos que delas decorrem ou a análise das formas de consumo. As selecções programáticas não resultam apenas das preferências e expectativas dos telespectadores, existindo uma grande diversidade de factores condicionantes, que podem favorecer a selecção ou a rejeição de determinados conteúdos. Em função dos contextos poderá existir uma maior ou menor actividade e utilidade subjacente às decisões de acesso ao *media* e dos conteúdos seleccionados. A descrição das diferentes perspectivas de análise existentes neste âmbito e o estado da arte consta respectivamente dos *Apêndice D* e *Capítulo II*.

Podemos identificar inúmeras disciplinas implicadas nas diversas vertentes de análise do sector televisivo e *Huertas (2002)*, também, salienta que as Ciências da Comunicação, têm-se desenvolvido desde diferentes perspectivas, basicamente desde a Sociologia, Psicologia, História, Antropologia, Literatura, Linguística e Política. A autora refere que os maiores progressos produziram-se desde a Economia e a Psicologia, uma vez que compreendemos melhor o funcionamento do mercado e do cérebro, do que da sociedade em si.

Por exemplo, de acordo com *Cassetti e Chio (1997)*, a Economia intervém na análise das empresas do sector televisivo, na análise de mercado dos seus produtos, no estudo do sector audiovisual e dos seus recursos, entre outras temáticas. A Sociologia visa o estudo de aspectos estruturais do sistema televisivo e de um modo geral a análise da oferta e do consumo. A Microsociologia centra-se especificamente nas diferentes formas de ver TV e nas dinâmicas sociais que se originam em determinado contexto. A Etnologia observa e descreve as situações e as dinâmicas que se geram em torno à televisão. A Psicologia analisa a natureza das relações que se estabelecem entre a televisão e os indivíduos ao nível das percepções, motivações, atitudes e comportamentos. A Pedagogia analisa a funcionalidade da televisão enquanto instrumento educativo e finalmente a Semiótica analisa a televisão enquanto meio de comunicação, centrando-se nos processos de significação; em todas as formas dos discursos, nas linguagens utilizadas e nos conteúdos transmitidos.

Cada disciplina aborda questões diferenciadas, existindo, também, divergências quanto aos métodos e fontes de dados adoptadas. Relativamente aos instrumentos ou técnicas de análise podemos recorrer aos métodos que registam a exposição, como por exemplo, à audimetria e aos diários de consumo; aos métodos que interrogam o telespectador, ou seja, às entrevistas e ao inquérito, mas, também, a métodos de observação, a análises de conteúdo e a modelos interpretativos, entre outras. As ciências Matemáticas contribuem com os seus próprios instrumentos, condicionando fortemente a informação que pode ser extraída das bases de dados constituídas através dos diversos instrumentos de que dispõe o investigador.

Nomeadamente a Estatística, entre outras áreas de análise de dados, contribui fundamentalmente como os seus próprios instrumentos e âmbito para a realização das análises quantitativas propostas e que vão permitir caracterizar a evolução dos indicadores de audimetria e o desenvolvimento de modelos explicativos e de previsão, assim como, dos comportamentos de exposição televisiva e selecção programática. Nas análises a desenvolver os métodos quantitativos destacam-se fortemente, correspondendo o principal objectivo desta tese à avaliação das potencialidades e do contributo de áreas metodológicas mais recentes para alguns dos contextos de análise que condicionam fortemente a gestão deste sector.

Para além da grande diversidade de disciplinas implicadas na análise do sector televisivo, também, podemos diferenciar as abordagens existentes quanto aos âmbitos (investigação teórica e prática), instrumentos e métodos de análise (investigação qualitativa e quantitativa) e relativamente às áreas de investigação, como por exemplo, a análise das entidades e da estrutura da audiência; a análise das atitudes sobre o consumo ou da satisfação; o estudo das motivações do consumo; a análise das reacções imediatas; a análise multivariada dos dados; a análise das dinâmicas e dos comportamentos associados ao consumo; a análise do conteúdo transmitido ou dos textos televisivos e a análise do *media* e dos processos culturais (Cassetti & Chio, 1997).

Perante a diversidade de disciplinas utilizadas no estudo do sector televisivo, Huertas (2002) refere que se deve deixar de observar os fragmentos disponíveis do conhecimento como peças irreconciliáveis para passar a entendê-los como elementos complementares entre si. Por exemplo, segundo a autora os âmbitos das Humanidades (Literatura, Linguística, Semiótica, História, Artes Sociais) e das Ciências Sociais (Sociologia, Política - Economia, Psicologia, Antropologia) deveriam trabalhar em conjunto. A ruptura que se verifica, por exemplo, entre as Humanidades e as Ciências Sociais, estende-se também às “Ciências Naturais e Formais”. Verificamos que as Ciências Estatísticas associadas às “Ciências Formais”, por vezes, não são valorizadas por parte de outras Ciências, que promovem em maior grau a investigação mais qualitativa (Cassetti & Chio, 1997; Fonseca, 2008; King, Tomz & Wittenberg, 2000). Segundo Huertas (2002), impera a falta de objectividade e múltiplas perspectivas são válidas.

No futuro, é possível que se verifique a fusão das Ciências Sociais com as Ciências Naturais, depois do processo de separação vivido no passado. Nesse sentido, já estão a emergir disciplinas fronteiriças, como a genética do comportamento ou a Sociobiologia, que estudam a base hereditária da conduta, e já comprovaram como esta interfere em questões como a formação do gosto pela beleza ou a paisagem.

Consequentemente, os resultados e as conclusões apresentadas, por vezes, encontram-se influenciadas pelas fontes de dados e métodos analíticos utilizados, condicionando a perspectiva segundo a qual se analisa o sector televisivo. As metodologias devem ser função dos objectivos da investigação, e não resultar do que mais convém e se adequa à formação do investigador (*Huertas, 2002*).

C.2. AS FONTES DE DADOS UTILIZADAS PARA A ANÁLISE DA EXPOSIÇÃO TELEVISIVA

Com base na bibliografia disponível nas principais revistas internacionais da especialidade da Comunicação Social associada ao sector televisivo (“Journal of Broadcasting & Electronic Media”, “Journal of Advertising”, “Journal of Advertising Research”, “Admap”, “Journal of Marketing”, entre outras) verificamos que as temáticas abordadas são inúmeras e muito diversificadas, abrangendo desde trabalhos de cariz mais qualitativo que visam a caracterização de aspectos relacionados com o consumo e os conteúdos programáticos, baseando-se em “Pretests” ou na observação de grupos de controlo, até à realização de análises quantitativas que permitam caracterizar, explicar e em última análise prever a evolução dos padrões de consumo da população ou de determinados índices de audiências relevantes para os profissionais do sector.

Os trabalhos quantitativos recorrem aos inquéritos e aos instrumentos de medição das audiências televisivas. Os principais métodos de medição das audiências são a audimetria, *audio-matching* e os diários de audiências que atendendo à crescente complexidade dos comportamentos de consumo televisivo permite também uma caracterização da exposição televisiva mais fiável e objectiva.

C.2.1. Os instrumentos de registro: a audimetria e os “Ratings”

As análises quantitativas baseiam-se nos diferentes métodos de caracterização e medição da exposição televisiva, como, por exemplo, em bases de dados obtidas pelo consolidado e mundialmente generalizado processo de audimetria. A investigação quantitativa encontra-se associada em maior grau com o Continente Norte Americano (*Huertas, 2002*), existindo um vasto número de publicações académicas no âmbito da audimetria, uma vez que as fontes alternativas de dados apresentam maiores desvantagens para a caracterização da exposição televisiva. Os Estados Unidos desempenharam um papel fundamental na área da medição das audiências televisivas, através do desenvolvimento da RDD (“Random – Digital – Dial”), assim como dos diários e do audímetro. A audimetria corresponde ao método que apresenta maiores vantagens. Como prova da falta de validade e fiabilidade dos resultados obtidos quando se aplicam os diários de consumo, existe um estudo realizado na Austrália em 1989. Demonstrou-se que os resultados dos diários subestimam o consumo televisivo em 20% a 33% (*Huertas, 2002*).

A nível Europeu, os reduzidos esforços desenvolvidos nesse sentido, acabam por ficar restringidos ao meio empresarial (*Weber, 2002*) e imperam os trabalhos recorrendo aos inquéritos (*Huertas, 2002*). Nos Países Anglo-Saxónicos, também, se utilizam outras fontes quantitativas de dados, como por exemplo, os diários de audiências e a audimetria quando se pretende caracterizar a exposição televisiva. Apesar de alguns autores poderem considerar que os sistemas de informação apresentam algumas fragilidades, proporcionando dados poucos válidos e fiáveis, também encontramos referências bibliográficas com pontos de vista contraditórios. Por exemplo, *Zarate (1997)*, refere que a audimetria permite aos investigadores dispor pela primeira vez, nas Ciências Sociais, de um instrumento metodológico de valor incalculável, para realizar todo o tipo de experiências e análises de dados, que em qualquer outro campo seria inviável por motivos económicos. Salienta que viabiliza a validação de quase trinta anos de prática profissional sem que tenham existido muitas vezes críticas à mesma. De acordo com a autora, os inquéritos correspondem a um pobre e rudimentar método de recolha de dados.

Infelizmente, impera o desconhecimento relativamente às potencialidades e características desta tecnologia. Contudo, a nível internacional e em especial nos Estados Unidos e em Inglaterra é frequente a utilização da audimetria para caracterizar comportamentos de consumo televisivo, constando uma síntese dos trabalhos mais relevantes, atendendo à temática em análise no *Capítulo III*.

Esta tese pretende demonstrar as potencialidades de determinadas área de análise estatística mais recentes para a resolução de questões que preocupam o sector televisivo, ou seja, a previsão das audiências televisivas, assim como, para a caracterização e o desenvolvimento de modelos explicativos do consumo televisivo, recorrendo aos sistemas de informação.

A tecnologia nacional não permite que dados sobre opiniões dos telespectadores sejam recolhidos, o que poderá no âmbito da investigação mais qualitativa ser entretanto facilmente resolvido com a realização de um inquérito ao painel ou recorrendo a outras abordagens que permitam esclarecer os aspectos que a audimetria não viabiliza, e, assim, complementar, isto é, não substituir as restantes análises que podem ser realizadas. A audimetria é uma ferramenta de investigação de mercados aplicada à investigação de audiências, adoptada em mais de oitenta países e aceite por todos os agentes implicados; que utiliza uma metodologia consensual e com fortes controles de qualidade (internos e externos) e que representa, também, uma via de conhecimento do comportamento da sociedade relativamente ao meio televisivo uma vez que disponibiliza uma grande quantidade de dados que caracteriza e quantifica a exposição televisiva (*Jauset, 2008*).

A audimetria corresponde a um sistema de informação que viabiliza a medição e a caracterização de diferentes aspectos relacionados com a exposição da população ao meio televisivo, contribuindo essencialmente para a investigação quantitativa dos meios de comunicação social, e existindo um vasto número de publicações internacionais que recorrem a esta tipologia de dados.

Os audímetros adoptados, actualmente, na maioria dos países registam comportamentos e não opiniões. Contudo, em 1988, por exemplo, na Suíça, já se realizavam sondagens qualitativas através do audímetro interactivo¹. No televisor surgem perguntas que o telespectador pode responder através de um comando à distância que se encontra incorporado no audímetro.

Todos os métodos de medição de audiências apresentam limitações e a audimetria não constitui uma excepção. Apesar das suas limitações, inúmeros especialistas definem-na como a melhor ferramenta de entre as conhecidas, apresentando grandes vantagens face aos restantes métodos de medição de audiências existentes, como se demonstra no *Apêndice A.I* (Jauset, 2008). Desde o seu aparecimento, em 1936, tem sido continuamente aperfeiçoado, tendo a empresa Nielsen testado mais de 200 modelos entre os anos cinquenta e o fim da década de oitenta, dos quais apenas 24 resultaram operacionais (Martin, 1989).

Têm sido observados contínuos avanços ao nível da tecnologia de medição das audiências sendo esta a ferramenta por excelência no futuro, quer para investigação quantitativa, quer para complementar a qualitativa. Por exemplo, os audímetros passivos, por vezes, substituem o comando por um sensor que detecta a presença dos telespectadores; aos audímetros podem ser associados códigos de barras e existem audímetros em forma de almofadas ou em pulseiras e que por vezes apresentam detectores fotoeléctricos.

Os métodos de medição de audiências contínuos e permanentes, uma vez que realizam uma análise diacrónica da audiência, reafirmam-se como os instrumentos de previsão por excelência, permitindo que modelos do comportamento sejam estabelecidos com a maior precisão possível. Os métodos de sondagem de carácter pontual que apenas se realizam uma vez, apenas permitem conhecer a audiência pelo horizonte temporal a que corresponde o estudo, oferecendo uma visão estática (sincrónica) do público o que impede a realização de previsões, uma vez que não reflecte uma evolução ou tendência.

¹ A descrição do audímetro interactivo consta em “La télévision Suisse” em Euroaudience nº14, Outubro de 1988, pp. 8-16.

Inúmeros são os campos de aplicação desta tipologia de dados, entre os quais se destacam: (1) a análise da eficiência das estratégias de programação; (2) a avaliação do desempenho dos conteúdos transmitidos; (3) a avaliação da capacidade explicativa e/ou predictiva das forças estruturais da indústria. Por exemplo, os modelos explicativos e/ou de previsão da evolução dos “ratings” considerando apenas a repercussão de factores estruturais. O “rating” enquanto tipologia de dados obtidos pelo processo de audimetria ou como o indicador de quantificação da exposição ao meio televisivo e principal variável que condiciona a construção dos planos de media e (4) a análise de padrões de consumo televisivo a nível programático e comercial, por exemplo, a avaliação da possível existência de hábitos de consumo televisivo, grau de regularidade da exposição; o comportamento da audiência no decorrer dos espaços comerciais, entre outras vertentes de análise.

C.2.2. Os inquéritos e as entrevistas

As desvantagens de inúmeras fontes de dados, como os inquéritos e as entrevistas, entre outras, vão muito mais além das afirmações que constam em *Fernandes (2001)*. Por exemplo, segundo *Mandinaveitia (1989)* ou *Wert (1990)*, entre outros autores, em inquéritos a amostra não é representativa uma vez que, em geral, resulta em maior grau de métodos de amostragem não probabilísticos, por ser mais rápido e menos oneroso de serem realizados e atendendo a que as pessoas que respondem estão mais interessadas no assunto abordado do que as pessoas que não respondem.

Também, salientam muitas outras desvantagens, entre as quais, as seguintes: a pessoa inquirida pode negar-se a colaborar voluntariamente e intencionalmente; o inquirido pode colaborar através da mentira e nem sempre é fácil a sua identificação; o inquirido pode errar involuntariamente, como consequência de falta de memória e de excesso de fidelidade a um canal ou programa; as respostas incorrectas podem resultar do “efeito hábito” uma vez que por estar familiarizado com determinado programa podem responder que viram a última edição mesmo que não o tenham realizado; o canal e espaço através do qual se realiza uma entrevista influencia os resultados, assim como o estilo, atitude e linguagem verbal do entrevistador (segundo a Kinésica até a forma de vestir do entrevistador pode deturpar os

resultados); apresenta a forte desvantagem de depender da memória de cada inquirido, sendo fornecidas respostas incorrectas por esquecimento ou por efeito “confusão”, e, assim, sucessivamente.

A entrevista corresponde a uma situação difícil, já que é uma forma de interacção social na qual os interlocutores se relacionam involuntariamente e nem sempre estão suficientemente motivados. Logo, é uma situação imposta para ambos os intervenientes, o que se pode repercutir negativamente no seu desenvolvimento e nas conclusões extraídas.

C.3. OS MÉTODOS DE INVESTIGAÇÃO APLICADOS AOS ESTUDOS DAS AUDIÊNCIAS

Relativamente aos métodos de investigação aplicados ao estudo dos *media* e das suas audiências, podemos diferencia-los em função de dois critérios: a finalidade da investigação, existindo neste âmbito os estudos exploratórios, descritivos e os causais, e as metodologias utilizadas que podem ser diferenciadas em quantitativas e qualitativas.

C.3.1. A finalidade da investigação

Relativamente à finalidade da investigação, os trabalhos podem ser classificados em: pesquisas exploratórias, descritivas e explicativas. A pesquisa exploratória tem como objectivo desenvolver, esclarecer e modificar conceitos e ideias, com base na formulação do problema ou hipóteses para estudos posteriores. A pesquisa descritiva pretende descrever as características de determinada população, fenómenos ou o estabelecimento de relações entre as variáveis. A pesquisa explicativa caracteriza-se pela identificação das causas que contribuem para a ocorrência do fenómeno.

As análises a realizar enquadram-se maioritariamente no âmbito da investigação prática, quantitativa, exploratória e descritiva.

Correspondem a análises exploratórias uma vez que se promove a identificação das variáveis estruturais de audimetria mais relevantes e respectivas interdependências, assim como a avaliação de possíveis alternativas de abordagens estatísticas que permitam explicar e prever a evolução dos principais indicadores de audimetria que condicionam a sustentabilidade das organizações. Simultaneamente, também, se refere as potencialidades de determinadas áreas metodológicas para a caracterização do consumo televisivo e para a avaliação de como os actores sociais interagem como este *media*.

Correspondem a análises descritivas, uma vez que se caracteriza a exposição televisiva, sob várias vertentes, recorrendo, sucessivamente, a diferentes tipologias de dados que permitem ir decompondo as análises e explicando a evolução dos principais indicadores de audimetria. Sendo o mercado televisivo muito volátil e vulnerável a alterações verificadas a diversos níveis, pretende-se demonstrar a necessidade de desenvolver análises explicativas e de previsão de audiências programáticas e comerciais, recorrendo a bases de dados de audimetria e a metodologias de análise estatística linear e não linear. Atendendo à dimensão das bases de dados, as metodologias a adoptar não podem ficar restringidas aos métodos mais tradicionais de análise, sendo necessário recorrer, também, ao *Data Mining* (Branco, 2010) e à análise Simbólica (Bock & Diday, 2000). As análises de previsão podem ser realizadas considerando apenas o factor temporal e com base em análises univariadas ou considerando a repercussão de variáveis exógenas em análises multivariadas, que quando não são na sua totalidade estimáveis ou conhecidas no momento de realizar uma previsão, adquirem a designação de análises explicativas segundo a bibliografia da especialidade (Weber, 2002).

Verifica-se, também, uma forte diversidade de contextos explicativos e de previsão que podem incidir em indicadores de audiências ou na quantificação do número de telespectadores que apresentam exposição a determinado suporte. Do ponto de vista do planeamento de *media* apresenta uma maior relevância a estimação de alguns dos indicadores, como, por exemplo, os *ratings*. Em determinados contextos a previsão, também implica a prévia caracterização de como a população interage maioritariamente com o meio televisivo, a nível programático e comercial.

Um sistema de previsão, também, pode promover a caracterização dos hábitos de audiência de cada programa onde se inserem os blocos de anúncios ou as inserções publicitárias, o que pressupõe para um planificador conhecer a duplicação e “replicação” que representa a duplicação das audiências face ao mesmo suporte, assim como a estrutura das audiências, o seu perfil.

C.3.2. A metodologia quantitativa e qualitativa

Em relação aos métodos utilizados podemos diferenciar entre “métodos qualitativos” orientados para o conhecimento e compreensão de atitudes, opiniões e motivações e os “métodos quantitativos” aplicados com o intuito de conhecer um tipo de informação quantitativa em forma de Estatística (*Espinosa, 2006*). As metodologias qualitativas são geralmente consideradas como pesquisa presente ou de acção ou até mesmo por levantamentos feitos através de questionários abertos, gravações, declarações, análises de grupo, entre outras abordagens.

Como foi referido no capítulo de introdução, os contextos sócio-económicos-culturais e de estruturação do mercado televisivo, condicionados pelos âmbitos geográficos e temporais considerados, influenciam fortemente os resultados, podendo levar a conclusões contraditórias, assim como consequência da disparidade existente quanto à formação académica dos investigadores e métodos de análise. A existência de duas perspectivas antagónicas associadas à análise da exposição televisiva resulta do leque muito abrangente de disciplinas que concorrem para a sua análise, e que nomeadamente se pode classificar em dois grandes segmentos: a investigação qualitativa e a quantitativa das audiências televisivas.

Existe nomeadamente também uma necessidade de combinar as possibilidades do quantitativo e do qualitativo, uma vez que a melhor via para a análise deste sector no futuro será a da análise interdisciplinar (*Huertas, 2002*). Cada parceiro da investigação deve contribuir com a aplicação dos seus métodos, disponibilizando aos restantes os resultados, ideias e comentários, numa tentativa combinada de conseguir extrair um melhor conhecimento sobre os fenómenos em análise. (*Fonseca, 2008*)

Verificamos que esta divisão de perspectivas, entre quantitativo versus qualitativo, bem notória na bibliografia da especialidade, não tem contribuído para que neste sector de actividade se promova um conhecimento mais claro de como a população interage com este *media*. A oposição entre o qualitativo e quantitativo corresponde a uma oposição clássica da tradição epistemológica, que indica também divisões académicas e distinções no tipo de organizações. Inúmeros centros de investigação estruturam-se precisamente em duas secções: qualitativa e quantitativa.

A grande maioria dos trabalhos quantitativos publicados no âmbito da audimetria nas principais revistas da especialidade referem-se a análises realizadas em universidades Norte Americanas, Britânicas ou em outros países anglo-saxónicos e em menor grau ao nível Europeu. A nível Europeu perante esta lacuna bibliográfica, por vezes realizam-se estudos comparativos face ao que se verifica, por exemplo, nos Estados Unidos.

A presente tese, também, visa contribuir para suprimir a lacuna existente. Em *Huertas (2002)* refere-se a existência de divergências quanto aos métodos de investigação na área da comunicação de massas entre os dois continentes. Segundo a autora, quando se classifica de forma global as abordagens existentes nesta área, será pertinente recorrer ao seu lugar de procedência, distinguindo-se entre os Estados Unidos e a Europa. Como resultado das diferenças contextuais, históricas, económicas, políticas e sociais, o estudo da comunicação social de massas difere entre ambos os lados do Atlântico. Estabeleceram-se directrizes divergentes e inclusive contraditórias em algumas ocasiões, determinadas pelo país de origem, devido à diversidade de teorias e das ferramentas de análise. O confronto que se verifica entre o quantitativo, mais estendido na América do Norte e o qualitativo com uma maior implantação na Europa e América Latina exemplifica esta questão.

Relativamente ao nível da maior ou menor adequabilidade das abordagens metodológicas quantitativas versus qualitativas para a análise do consumo televisivo em *Huertas (1997, 2002)*, *Souchon (1993)*, *Saperas (1998)* e *Igartua e Badillo (2003)* podemos encontrar uma excelente exposição sobre esta temática.

Para *Huertas (2002)* é conveniente ressaltar que não existe objectividade, nem neutralidade quando se trata do estudo de aspectos sociais e culturais e que, consequentemente, a nossa visão dos factos depende do método com o qual o investigador se aproxima da realidade. Salienta que não existe uma visão mais correcta do que outra e a prova está em que nenhuma prevaleceu sobre as restantes. A eleição de uma ou outra depende de quem se aproxima da audiência e para quê. Um investigador, com uns princípios e seguindo uns passos metodológicos, obterá uma informação que complementar, negará ou modificará as visões de outros analistas com bases e caminhos metodológicos diferentes. E isto será o que permitirá o progresso do conhecimento, sempre e quando não fique num estéril confronto entre intelectuais e/ou académicos apenas por conseguir o domínio do seu território. Segundo a autora o problema é que apesar da grande quantidade de estudos empíricos que se têm vindo a acumular no decorrer do século XX não existe uma recopilação sistemática que solucione a fragmentação do conhecimento das audiências.

Nesse sentido, na presente tese, apresentam-se os potenciais de aplicação de áreas metodológicas mais recentes, não só para o desenvolvimento de modelos de “*ratings*”, mas, também, para a análise da interacção dos telespectadores com o meio televisivo. Segundo a autora, o quantitativo contribui com uns dados finais e o qualitativo oferece um tipo de informação que ajuda a compreender o estudado ao contribuir com argumentos que explicam o que no primeiro caso aparece reflectido por valores mensuráveis, absolutos ou percentuais (*Huertas, 2002*). Portanto, as abordagens metodológicas que visam extrair o máximo de informação das bases de dados de audimetria contribuem para a identificação de situações, factos, comportamentos que posteriormente podem ser desenvolvidos pelas áreas que têm essa responsabilidade.

C.3.3. As abordagens analíticas

A investigação realizada nas Ciências Sociais em geral e no âmbito dos *Media* em particular encontra-se penalizada por ser desvalorizado e preterido o recurso a metodologias de análise Estatística de dados de maior complexidade, e consequentemente, também, das fontes de dados que se encontram fortemente dependentes da sua aplicação (*Weber, 2002*).

Segundo *Zarate (1997)*, a utilização sistemática das Matemáticas e da Estatística nas Ciências Sociais é recente, tendo sido inaugurado em 1954 por *Lazersfeld* em “Mathematical Thinking in the Social Sciences”. Na área da Comunicação Social, as abordagens metodológicas que imperam na bibliografia da especialidade não viabilizam a análise dos actuais padrões de comportamento de consumo televisivo, assim como, dos índices de audiências, em especial em contextos de previsão ou explicativos da sua evolução (*Weber, 2002*). Em *Fonseca (2008)*, também consta uma exposição neste âmbito, referindo o autor a frequente desconfiança e rejeição da investigação quantitativa, assim como da informação numérica que resulta, segundo vários autores, do desconhecimento existente sobre a investigação quantitativa contemporânea.

Com base na análise bibliográfica realizada, facilmente, se constata que os resultados dos trabalhos desenvolvidos encontram-se penalizados por imperar as abordagens Estatísticas mais tradicionais. Foi possível verificar que as revistas da especialidade requerem que a descrição metodológica seja o mais sucinta possível e que os trabalhos que promovam uma utilização mais intensiva de metodologias estatísticas menos tradicionais acabam por ser considerados como sendo pouco adequadas ao público-alvo das mesmas, encontrando-se, inevitavelmente, estes trabalhos publicados em revistas de Estatística Aplicada ou de Investigação Operacional.

Podemos identificar publicações recentes no sector da Comunicação Social norte americana, por exemplo, para esclarecer que probabilidades condicionadas não são equivalentes a probabilidades conjuntas e que, consequentemente não produzem os mesmos resultados quando consideradas como variáveis dependentes (*Webster, 2006*) em análises de Regressão linear tradicionais: “*Eastman, Newton, Riggs & Neal-Lunsford (1997)* concluded that the potency of inheritance effects seemed to be in decline. However, that conclusion seems unwarranted because Eastman et. al were essentially comparing “apples and oranges.” (*Webster, 2006, p.327*). Durante décadas são apresentadas nestas áreas conclusões contraditórias, que resultam em maior grau da disparidade existente ao nível das bases de dados e abordagens metodológicas seleccionadas.

Consequentemente, a área da Estatística que viabiliza e serve de suporte a um vasto leque de análises realizadas neste sector condiciona, não só a previsão de indicadores fundamentais, como, também, uma melhor caracterização e percepção dos processos subjacentes ao consumo e à exposição televisiva. Na presente tese, procurou-se fazer um levantamento dos mais recentes desenvolvimentos em diversas áreas de análise estatística de dados, apresentando os respectivos potenciais de aplicação para a análise de bases de dados de audimetria e as correspondentes mais valias.

De acordo com *Fonseca (2008)*, verificamos que a maioria dos trabalhos realizados nas Ciências Sociais utilizam algumas metodologias multivariadas e adoptam como fonte principal de dados quantitativos, os inquéritos. Nesse sentido, esta tese procura contribuir com a avaliação das potencialidades de áreas metodológicas mais recentes para o desenvolvimento de análises explicativas e de previsão de índices de audiências, no entanto também constam algumas abordagens analíticas que podem promover uma melhor caracterização de como a audiência interage como este *media*.

As análises a realizar visam que certas questões apenas exploradas segundo determinados âmbitos Estatísticos, sejam aprofundadas com o recurso a abordagens Estatísticas complementares, comparando-se os resultados obtidos, por exemplo, promovendo a aplicação de análises não lineares onde imperam as lineares na literatura da especialidade internacional.

Ao contrário de *Fernandes (2001)*, que defende que a hipótese de uma investigação qualitativa da comunicação parece emergir como resultado da progressiva mudança dos processos de consumo, segundo *Cooper e Tang (2009)*, inclusive os inquéritos não viabilizam uma análise que possa ser considerada clara da exposição televisiva. Apenas a utilização de instrumentos que caracterizem com uma maior profundidade e precisão o consumo televisivo pode viabilizar um melhor conhecimento de como a audiência interage com este *media*, entre os quais se encontra a audimetria que apresenta inúmeras vantagens face os diários de audiências. O aperfeiçoamento da tecnologia para registar comportamentos, mas também aspectos mais qualitativos da exposição televisiva será o que viabilizará no futuro um melhor conhecimento e caracterização do consumo televisivo.

A extracção de conhecimento útil encontra-se, por sua vez fortemente dependente da evolução observada nos instrumentos metodológicos. Por exemplo, inesperados desenvolvimentos ao nível do grau de previsibilidade das trajectórias geográficas dos indivíduos levadas a cabo pela equipa do Físico Prof. Barabási (director do Northeastern's world - leading Center for Complex Network Research) não seriam viáveis sem os sistemas de informação e as metodologias associadas ao *Data Mining* e não só. Verificamos que após décadas de investigação recorrendo aos métodos analíticos e fontes de dados mais generalizadas, continuam a persistir inúmeras incógnitas e contradições subjacentes à exposição televisiva.

A informação contida nas bases de dados de audimetria, infelizmente, também não tem sido extraída, como consequência das limitações das ferramentas estatísticas existentes e da dificuldade de acesso aos dados em geral detidos por multinacionais, assim como a nível internacional, por vezes de uma incorrecta codificação e interpretação dos dados ou de serem adoptados métodos pouco adequados à análise de bases de dados que apresentam as suas especificidades. No entanto, nos últimos anos verificou-se uma significativa evolução neste âmbito, como o desenvolvimento da Análise Simbólica e ao nível das inúmeras áreas do *Data Mining*, o que viabilizará que esta realidade seja alterada. Para além de demonstrar as potencialidades da audimetria e de realizar uma análise metodológica que permita a identificação das abordagens mais adequadas a determinados objectivos, também, procura o desenvolvimento de análises descritivas, explicativas e de previsão dos índices de audiências, recorrendo a metodologias que em geral não constam da bibliografia internacional, nomeadamente as análises não lineares e as simbólicas.

Também neste sector, assim como *Fonseca (2008)* salienta, imperam as abordagens lineares, recorrendo à Regressão linear tradicional. Segundo alguns autores, são escassos os trabalhos quantitativos realizados no âmbito das análises explicativas e de previsão de índices de audiências, que promovam a utilização de abordagens estatísticas mais recentes, e os que existem não se encontram acessíveis aos investigadores académicos uma vez que os resultados da investigação realizada acabam por não ser divulgada pelo valor comercial que apresentam (*Weber, 2002; Danaher, Dagger & Smith, 2011*).

C.4. A INVESTIGAÇÃO ACADÉMICA E A EMPRESARIAL

Apesar das bases de audimetria serem, frequentemente utilizadas para a análise da exposição televisiva a nível académico, uma segunda questão a ser abordada terá que ser a relação existente entre a investigação comercial e a académica. Apesar desta tese não privilegiar no global a análise do consumo televisivo em si, mas, sim, de índices de audiências apresentam-se algumas considerações a este respeito. Relativamente à validade dos dados de audimetria e aos dois âmbitos de investigação existentes, em *Huertas (2002)* podemos encontrar uma excelente exposição sobre esta temática. A autora entende que será imprescindível sublinhar o contributo de um enfoque multidisciplinar. Depois de décadas de esforços por forma a encontrar soluções de um ou do outro lado, a renovação socioeconómica e tecnológica faz com que a disjunção entre um conhecimento académico e um conhecimento “administrado” das audiências se encontre neste momento mais aberto do que nunca.

Ford, Pitt e Okazaki (2008) editores de uma das principais revistas internacionais da especialidade, o *Journal of Advertising Research*, salientam a necessidade de existir uma maior cooperação entre os autores por forma a tirar partido de possíveis sinergias, e quanto ao desenvolvimento de temáticas globais, uma vez que tem existido nas últimas décadas uma maior incidência em trabalhos que focam aspectos intrínsecos a determinados mercados e consequentemente não generalizáveis. Os autores, também, salientam a necessidade dos trabalhos realizados começarem a apresentar uma maior relevância prática, ao invés de serem privilegiados trabalhos excessivamente teóricos que em nada contribuem para um melhor desempenho organizacional.

Na presente tese, privilegiou-se uma temática que apresenta uma grande relevância prática, sendo determinante para o bom desempenho organizacional. Segundo *Cassetti e Chio (1997)* a orientação de uma investigação também depende de factores bastante concretos, como os âmbitos. No entanto, apesar de manterem as suas principais diferenças, actualmente, tendem a difuminar os seus respectivos campos de aplicação.

Em concreto, as grandes empresas conjugam com frequência as perspectivas teóricas e práticas, enquanto que as perspectivas académicas já recorrem frequentemente a uma ampla e sistemática recolha de dados empíricos práticos. A nível nacional, também, têm existido parcerias entre o meio empresarial e académico no âmbito da análise das bases de audimetria. No *capítulo II* encontra-se uma síntese das variadas publicações académicas que promovem a análise de bases de dados de audimetria.

C.5. ENQUADRAMENTO DAS ANÁLISES A REALIZAR

No *Quadro C.5.1* consta uma síntese dos âmbitos de análise no sector televisivo. As análises a realizar abordam o desenvolvimento de modelos explicativos e de previsão das audiências. Uma vez que as previsões exactas não são viáveis nas Ciências Sociais e em geral as abordagens conduzem a margens de erro que podem ser muito significativas, é necessário promover, também, as análises explicativas e identificar as abordagens analíticas que o possam viabilizar com as menores margens de erro possíveis.

Nenhum sistema de previsão será exacto porque estamos a lidar com comportamentos sociais. Em análises de previsão será necessário ter em atenção a sua rapidez, ter conhecimento capitalizável, baixo custo e devem-se privilegiar as análises explicativas, ou seja, não prever apenas, mas, também, interpretar e evolução dos dados. As análises propostas não correspondem ao âmbito da Psicossociologia, entre outras áreas, que promovem o estudo das atitudes, interesses e das motivações subjacentes ao consumo televisivo ou da oferta televisiva; a análise das reacções dos telespectadores; a análise da dinâmica dos comportamentos de consumo; a análise dos *media* e dos processos culturais, a análise dos textos transmitidos e conteúdos transmitidos, entre outras temáticas, que apresentam um carácter mais qualitativo.

Na realidade foi realizado um maior investimento temporal a promover a identificação de uma grande diversidade de métodos analíticos que fossem os mais adequados à análise de bases de dados massivas de audimetria e na estruturação de bases de dados com uma elevada dimensão e que apresentam a informação demasiado dispersa e com uma estrutura que em nada se adequa à realização das análises propostas.

Por isso mesmo, na bibliografia da especialidade, após várias décadas impera a análise dos indicadores universais, referentes ao universo televisivo, apesar de alguns autores referirem que os dados já lhes foram disponibilizados de forma estruturada pelas empresas que disponibilizaram os dados e que nesta tese não se verificou.

Ainda assim, neste trabalho sempre que viável consideram-se as variáveis sociodemográficas que constam das bases de audimetria. Apenas está em causa contribuir para a resolução de questões que apresentam grande relevância prática, condicionando fortemente o desempenho organizacional de todos os agentes associados à indústria televisiva. Nesse sentido, promove-se o desenvolvimento de análises explicativas e de previsão de índices de audiências, avaliando o contributo de abordagens metodológicas associadas a áreas mais recentes. Existem publicações de editores das revistas da especialidade que criticam a falta de relevância prática dos trabalhos apresentados (*Ford, Pitt & Okazaki, 2008*). Pretende-se avaliar a repercussão de variáveis estruturais, entre as quais as associadas à estruturação da grelha de programação, na evolução dos principais índices de audiências promovendo simultaneamente a sua previsão. Esta tese, apresentando inúmeros potenciais de aplicação das áreas metodológicas mais recentes, proporciona um forte contributo ao nível dos instrumentos analíticos que viabilizam a caracterização de diferentes vertentes da exposição televisiva.

Com base nos resultados alcançados, algumas conclusões, também, são retiradas a respeito de como o painel interage com o meio televisivo. A motivação para a realização desta tese, surge das dificuldades sentidas no meio empresarial quando é necessário prever e caracterizar a evolução de determinados indicadores que caracterizam a exposição televisiva do painel de audimetria. Consequentemente, as análises realizadas que resultam de um trabalho individual visam contribuir para a futura previsão dos índices de audiências, incidindo na identificação dos métodos analíticos mais adequados à análise de bases de dados de audimetria em função de determinados contextos e na realização de análises explicativas dos índices de audiências.

Não se pretende apenas prever determinados indicadores com base na análise de séries temporais, ignorando, totalmente, a repercussão de factores associados ao planeamento de *media*, mas, sim, sempre que possível avaliar como as variáveis estruturais explicativas identificadas condicionam a sua evolução. As metodologias e as fontes de dados correspondem às necessárias e adequadas aos objectivos propostos, como se demonstra nos capítulos seguintes. As análises a realizar visam contribuir para a resolução de questões que são relevantes no âmbito do planeamento dos media.

As ciências Estatísticas, por sua vez, contribuem para o estudo do sector televisivo através dos seus próprios instrumentos que viabilizam a realização de análises quantitativas, fundamentais para a análise de bases de dados de audimetria ou de qualquer outra tecnologia de medição das audiências. Logo, não está e causa a análise das dimensões sócio-culturais e a análise das variáveis sociais e culturais, ao nível dos diferentes interesses.

Apesar destas abordagens se basearem na quantificação das audiências, esta tipologia de análise, também, inclui aspectos qualitativos, no entanto em poucas ocasiões este aspecto acaba por ser valorizado (*Huertas, 1997*). As variáveis sociais e culturais, nesta tese, foram utilizadas com o intuito de avaliar a sua repercussão na evolução e variabilidade observada nos indicadores de audimetria. Estando em causa promover a futura previsão das audiências, e assim como se verifica na literatura da especialidade que aborda a análise de indicadores de duplicação de audiências e dos restantes indicadores de audimetria, os valores para o universo televisivo foram os privilegiados nas análises a serem realizadas.

A presente tese, não tem como objectivo a apresentação de resultados que se encontram associados em maior grau à investigação mais qualitativa e que persegue objectivos diferentes. Apenas a investigação quantitativa foi realizada, com o objectivo de caracterizar os indicadores de audimetria, promovendo a sua previsão e a explicação da sua evolução, atendendo às influências estruturais que podem ser avaliadas com base na tipologia de dados que foram disponibilizados.

As análises realizadas enquadram-se nas teorias estruturais do consumo televisivo, que correspondem às que são necessárias para promover a previsão destes indicadores uma vez que apresentam maior poder de previsão e recorrem a dados de “*ratings*” para a análise de comportamentos de consumo televisivo a nível macro, não tendo em consideração os interesses individuais. Avaliam-se as influências estruturais nos índices de audiências, mas não se assume necessariamente a existência de um painel, essencialmente mais passivo nas suas selecções. Alguma investigação realizada com base na audimetria também já demonstrou a existência de alguma actividade. A audimetria, certamente, que não encara a audiência como passiva ou activa apenas regista os padrões de exposição televisiva, sejam eles quais forem.

Verificámos com base na análise de fluxos de audiências que existe alguma actividade, no sentido de transição entre canais, mas que acaba por apresentar uma menor repercussão na evolução dos *ratings* face às decisões de ligar e desligar o TV. *Cooper e Tang (2009)* com base em inquéritos realizados defendem a predominância das motivações ritualísticas e a influência das estruturas na exposição televisiva, assim como uma menor dependência dos resultados obtidos face às variáveis sociodemográficas, uma vez que apenas as faixas etárias apresentaram alguma relevância estatística quanto à diferenciação dos padrões de exposição.

Algumas características sociodemográficas encontram-se em maior grau relacionadas com os factores de disponibilidade da amostra do que com as selecções e exposição ao media. Verificamos que na literatura da especialidade, após várias décadas de investigação a nível mundial, não existe qualquer consenso relativamente a inúmeras questões, entre as quais a dicotomia actividade e passividade do público e quanto ao grau de heterogeneidade subjacente ao consumo televisivo ou em que medida as variáveis socioeconómicas, por si só condicionam e diferenciam os resultados (*Huertas, 1997*), sendo provavelmente necessário adoptar outras variáveis discriminantes do consumo televisivo em função do objectivo das análises

ÁREAS DE INVESTIGAÇÃO	ORIENTAÇÕES TEÓRICAS	ÂMBITOS
ANÁLISE DAS ENTIDADES E DA ESTRUTURA DA AUDIÊNCIA ESTUDO DAS ATITUDES ANÁLISE DA SATISFAÇÃO ANÁLISE DAS MOTIVAÇÕES ANÁLISE DAS REACÇÕES IMEDIATAS ANÁLISE MULTIVARIADA DE DADOS ESTILOS DE VIDA DINÂMICAS DE COMPORTAMENTOS ANÁLISE DE CONTEÚDOS ANÁLISE DOS TEXTOS TELEVISIVOS ANÁLISE DO MEDIA ANÁLISE DOS PROCESSOS CULTURAIS	O MEDIA O TEXTO O TELESPECTADOR O CONTEXTO A RECEPÇÃO OBJECTOS DE ESTUDO PRODUÇÃO Aspectos tecnológicos Aspectos económico-empresariais Aspectos culturais e sociais Aspectos político-institucionais OFERTA Programas Programação Mercado CONSUMO Índices de audiências Eleições de consumo Modalidades de visão Processos de Compreensão Processos de Valorização Efeitos Dieta de Media	INVESTIGAÇÃO TEÓRICA INVESTIGAÇÃO PRÁTICA INSTRUMENTOS REGISTRAR Audimetria Diários de Consumo Histórias Vitais OBSERVAR Observação Participante INTERROGAR Entrevista Conversação Inquérito Teste INVENTARIAR Análises de Conteúdos RESUMIR Esquemas de Análises Textuais RELACIONAR Categorias Modelos Interpretativos
FINALIDADE DA INVESTIGAÇÃO	DISCIPLINAS	METODOLOGIAS UTILIZADAS
EXPLORATÓRIOS DESCRITIVOS CAUSAIS	Economia Sociologia Psicologia Pedagogia Etnologia Estatística Semiótica Marketing Comunicação Social Direito Eng. Telecomunicações	QUANTITATIVAS QUALITATIVAS

Quadro C.5.1 – Um mapa da investigação na área televisiva
Fonte: Casseti & Chio (1997)

Uma vez que a nível internacional, as teorias estruturais correspondem às que apresentam uma maior capacidade de previsão será necessário identificar as abordagens metodológicas mais recentes que permitam maximizar as percentagens de variância explicada atendendo aos dados nacionais de “*ratings*” que servem de suporte internacionalmente a esta tipologia de análises, sendo, isso sim, o objecto de estudo desta tese. Contudo, como possível contributo para as Ciências Sociais, também, se enumeram as potencialidades das áreas metodológicas existentes que permitam uma caracterização da exposição individual, essencialmente ao nível do *Data Mining* Sequencial e Multirrelacional.

Do ponto de vista metodológico teve-se em consideração a apresentação das abordagens mais adequadas à análise de fluxos agregados de audiências, assim como a nível individual, o que permitirá contribuir para o esclarecimento das inúmeras incógnitas que existem, actualmente, viabilizando uma caracterização mais aprofundada dos comportamentos de consumo televisivo. Em *Fernandes (2001, p.133)* reconhece-se que: “A Audimetria permite a medição ao segundo, bem como uma análise detalhada do relacionamento dos telespectadores com a mensagem publicitária, captando cerca de 50 a 60 mudanças de canal por dia, ... Este sistema facilita uma avaliação do zapping entre canais.”

Quando possível, em função dos resultados das análises realizadas, apresentam-se, também, algumas considerações a respeito do consumo televisivo, para os segmentos sociodemográficos representados nas bases de dados de audimetria. Nomeadamente a avaliação da repercussão das variáveis socio-demográficas na evolução dos indicadores de audimetria, também, permite a caracterização da exposição televisiva para os segmentos considerados e a análise da duplicação de audiências conjugada com os restantes indicadores, permite a avaliação do grau de actividade versus passividade inerente ao consumo televisivo a nível nacional.

APÊNDICE D

OS MODELOS EXPLICATIVOS DO COMPORTAMENTO DO TELESPECTADOR

Neste apêndice, consta uma introdução à temática dos modelos explicativos da exposição televisiva, e que podem ser, maioritariamente, agregados em duas perspectivas teóricas de análise, que se confrontam desde há longa data, e que se diferenciam, fortemente, quanto aos pressupostos e métodos de investigação adoptados.

Devemos proporcionar uma visão mais geral das perspectivas teóricas que caracterizam o comportamento dos telespectadores, uma vez que nos permitem compreender o seu processo de decisão, beneficiando os profissionais do planeamento de *media* e as estações televisivas, com uma maior percepção sobre como devem agir para alcançar determinada reacção e exposição televisiva, contribuindo, assim, também, para a sua previsão.

As conclusões que resultam dos modelos explicativos do consumo televisivo condicionam a selecção das abordagens metodológicas e das perspectivas de análise que podem conduzir aos melhores resultados na área da previsão das audiências televisivas. Comparam-se as duas principais perspectivas de análise ao nível da sua capacidade predictiva da evolução dos fluxos agregados de audiências.

D.1. A ABORDAGEM TEÓRICA DO COMPORTAMENTO DO CONSUMIDOR

Os modelos de caracterização dos comportamentos dos consumidores têm como objectivo compreender o seu processo de decisão, analisando o maior número de factores, sendo uma área que se encontra em constante evolução, assim como os indivíduos.

Esta área beneficia os consumidores, as organizações e os profissionais do Marketing, uma vez que recorrendo a estes modelos podem agir de modo a conseguirem determinadas reacções por parte do seu mercado-alvo.

Nas análises devem ser consideradas variáveis externas ambientais e internas psíquicas que estão envolvidas nos processos, esquematizando as interdependências que podem ser observadas e que do ponto de vista teórico podem resultar na aquisição de determinado produto ou serviço. Os modelos são utilizados para compreender ou prever determinados comportamentos, eliminando os aspectos menos relevantes.

Analisa quem compra, porquê e como e, logo, contribuem para a previsão dos comportamentos que é decisiva para a avaliação das estratégias de Marketing que permitem corresponder às necessidades do seu público-alvo. Esta vertente de análise enquadra-se no Positivismo e que complementa a abordagem do Interpretativismo ou Experimentalismo, que procura os significados escondidos nos comportamentos de consumo.

D.2. AS PERSPECTIVAS TEÓRICAS QUE ORIENTAM A EXPLICAÇÃO DA EXPOSIÇÃO TELEVISIVA

Contrastando com a maioria da bibliografia existente no âmbito das análises de previsão das audiências televisivas, na presente tese, para a definição das abordagens estatísticas, foram consideradas as teorias explicativas que constam na literatura da Comunicação Social a respeito dos comportamento dos telespectadores.

Em geral as análises de previsão incidem, essencialmente, nos diversos índices de audiências, ignorando os comportamentos que condicionam a sua evolução. A forma como os telespectadores interagem maioritariamente com o meio televisivo e os programas condiciona as abordagens de previsão e explicativas que são mais adequadas em diferentes contextos. Quanto maior for a actividade do painel e a irregularidade da evolução dos indicadores a nível agregado, maior será a necessidade de compreender e caracterizar a evolução dos dados.

A investigação realizada neste sector conduziu à proliferação de modelos e concepções, inúmeras vezes contraditórios entre si, o que se deve à diversidade existente ao nível da formação académica dos investigadores, aos métodos de análise e dados utilizados (Webster, 2006), assim, como, também, aos contextos sócio-económicos e culturais dos países em tais estudos foram desenvolvidos, condicionando fortemente os resultados atingidos. A divisão da bibliografia em trabalhos qualitativos versus quantitativos, conduziu a duas abordagens e perspectivas antagónicas a respeito dos factores que condicionam a exposição televisiva.

Nos seguintes parágrafos consta um excerto de uma publicação, onde fica bem claro o confronto que se verifica entre a teoria das gratificações e usos do *media* e a teoria estrutural do consumo televisivo, assim como as dificuldades de integração das mesmas e que, prevalecem, segundo os autores, as influências estruturais (Cooper & Tang, 2009):

Scholars have long advocated empirical integration of active-audience and structural theories to best explain audience exposure to television. Two primary, though conflicting theoretical perspectives guided explanations of medium choice and use. The first emphasizes individual determinants most commonly articulated through the uses and gratifications approach. The second focuses on how structural characteristics such as audience availability, access and scheduling factors influence the size and composition of audiences. This study addresses the theoretical gap by integrating individual and structural factors that predict exposure to television. ...

Findings suggest that no single theoretical construct explains the complexities that determine exposure to TV. Exposure to television was influenced more by ritualistic motivations (unique explanation of 4.7%) and availability (3,6%) than by instrumental motivations (0,5%). This lends some credence to the idea that overall exposure to television is determined less by a “desire” for specific content than by factors of habit, availability, and access to the medium. This is not surprising given that ritualistic motivations for television are more habitual in concept and availability for television likely manifests itself in regular patterns, while instrumental motivations are more often associated to a desire for specific programs or program types. ...

Current audience theories continue to provide important insights, but considerable research is needed to fully explain how and why individuals use television (and other media). ... This study provides one step in this direction and encourages future inquiry into an ever-complex media world.

D.2.1. A perspectiva das gratificações e usos do *media*

A abordagem das gratificações e usos do *media* defende o carácter activo da audiência, sendo as suas escolhas função de preferências ou da satisfação de determinadas necessidades. A exposição encontra-se condicionada por características individuais e de interesse por determinados conteúdos, verificando-se uma exposição ou ausência de exposição de forma intencional. Contudo, o acesso ao *media* pode decorrer de hábitos desenvolvidos, existindo uma maior actividade apenas ao nível dos conteúdos seleccionados.

Esta abordagem encontra-se, essencialmente, associada às Ciências Sociais e ignoram a influência das estruturas na exposição ao *media*. “However studies in this tradition empirically ignore that exposure is not completely free of constraints.” (Cooper & Tang, 2009, p.2). Segundo estes autores, as motivações instrumentais, associadas a uma maior actividade e intencionalidade da audiência relacionam-se em maior grau com os eventos desportivos e os espaços informativos.

De acordo com Callejo (1995, p.252), do Centro de Investigação Sociológica CIS, a mudança na concepção do público decorre da evolução observada ao nível dos métodos de investigação: “Las necesidades de la investigación empirica son las que en mayor medida han expresado la exigencia del cambio en la concepción de la audiencia desde agente pasivo hasta agente activo.” Logo, a mudança de paradigma não decorre de um pensamento ideológico, mas, sim, da evolução “científica” ao nível das metodologias de investigação na comunicação mediática.

Segundo Marley (1992), a teoria das gratificações e usos do *media* corresponde a uma problemática essencialmente psicologista, porque se baseia em estados mentais, necessidades e processos abstraídos da situação social dos indivíduos. Segundo o autor grande parte da televisão é consumida como “televisão” e não diz respeito a um programa em particular.

Quando alguém se diz atraído por determinado programa, dificilmente resulta dos méritos de um episódio, mas, sim, de uma escolha de um exemplo de um género específico que satisfaz esse telespectador. Segundo *Shroder (1999)*, para além, da fragmentação do conhecimento ao nível dos *media*, a investigação qualitativa que se associa à teoria das gratificações e usos do *media* preocupa-se demasiado com a sua validade interna. Salienta, por exemplo, que os estudos culturais apresentam uma elevada validade interna, em detrimento da externa, o que limita as possibilidades de generalização e questiona a utilidade do trabalho realizado para conhecer a sociedade ou estabelecer uma teoria. Este desequilíbrio é criticado pela perspectiva exclusivamente teórica, sobretudo por aquela que não comparte a necessidade de observar um fenómeno social para conseguir desenvolver uma teoria.

D.2.2. A abordagem estrutural

A teoria dos factores estruturais explica melhor o comportamento da audiência quando esta assiste com passividade, aderindo às condicionantes estruturais e recorrendo à televisão esporadicamente para preencher os seus tempos livres ou vazios. Sob esta perspectiva será uma utilização ritualizada condicionada pela disponibilidade do público e influenciada em maior grau por parte dos factores estruturais, como, por exemplo, o acesso ao *media*, o hábito, os padrões de disponibilidade, os factores de planeamento e os comportamentos que decorrem da estruturação da grelha de programas. Os factores estruturais influenciam fortemente a dimensão e a composição da audiência, verificando-se uma exposição ou ausência de exposição não intencional. Esta perspectiva encontra-se, essencialmente, associada à Comunicação Social, em especial anglo-saxónica e à investigação quantitativa.

Os factores estruturais podem-se repartir por duas categorias: (1) Características de longo prazo do mercado televisivo e fora do controle dos canais, por exemplo, o nível de penetração da TV Cabo e (2) Os factores temporais e de planeamento televisivo, por exemplo, as variáveis de desempenho e caracterização de programas e espaços comerciais ou temporais e que são objecto de análise nesta dissertação a nível programático e comercial.

Um correcto planeamento da grelha de programas, também, pode contribuir para o desenvolvimento de determinados comportamentos, influenciando a evolução dos indicadores de audimetria, como, por exemplo, a duplicação de audiências entre programas, tornando complexa a relação existente entre as possíveis preferências do público e o que na realidade acaba por ser seleccionado.

Os trabalhos que visam a análise de índices de duplicação com base em dados de audimetria constituem uma grande fracção dos estudos internacionais desenvolvidos no âmbito dos “viewing choice models” e estão associados às teorias estruturais. Os comportamentos estruturais podem-se agrupar nas seguintes rubricas: (1) Lealdade a canais, que representa a tendência do público para seleccionar consecutivamente programas de determinado canal, (2) Efeitos de herança, que visam a análise de níveis de duplicação de audiências entre programas transmitidos no mesmo dia e canal e geralmente adjacentes, (3) Lealdade a programas que corresponde aos níveis de audiência duplicada entre diferentes episódios de um mesmo programa, geralmente consecutivos e por último (4) Exposições repetidas que se traduz no grau de exposição da audiência a retransmissões de um mesmo programa. A duplicação de audiências é imprescindível para a previsão das audiências televisivas e para o planeamento dos media (*Fernandes, 2001*), sendo, consequentemente, objecto de estudo da presente tese.

D.2.3. A comparação entre ambas as perspectivas

A teoria estrutural e a teoria de gratificações e usos do *media* diferenciam-se significativamente e perseguem diferentes objectivos e adoptam diferentes métodos de análise (*Cooper & Tang, 2009; Cooper, 1996*). As diferenças entre ambas as abordagens são muito significativas, constando uma síntese nos *Quadros D.2.3.1 e D.2.3.2*. Correspondem a duas perspectivas contraditórias, estando associadas a diferentes áreas científicas que se confrontam desde há longa data.

Vários autores, entre os quais *Morley (1998)*, salientam que o investigador-observador imparcial é uma figura impossível; o único que existe são realidades individuais que se vão gerando, logo não se verifica uma visão geral da realidade; a realidade não é fixa, e só pode ser abordada desde o cepticismo e o paradoxismo.

Lipovetsky (2000) aponta que a personalização (teoria das gratificações) e a homogeneização (teorias estruturais) não são fenómenos incompatíveis. Tendo em conta que a comunicação é social, destinada a um amplo grupo, por sua vez formado por diferentes agrupações e que a recepção é um processo individual, que perspectiva será a mais adequada? Segundo *McQuail (2000)*, quando alguém opta por estudar a audiência de uma determinada forma também está a decidir de que forma configurará a mesma. A escolha de uma ou outra depende em maior grau de quem se aproxima da audiência e para quê (*Huertas, 2002*).

Quando o objectivo corresponde à previsão das audiências, as teorias estruturais têm produzido os melhores resultados. Segundo determinados autores, as duas ópticas podem influenciar os padrões de consumo, defendendo a existência de alguma passividade quanto à decisão de ver televisão, mas, apresentando as selecções programáticas uma maior actividade. Para outro segmento de telespectadores a selecção, também, pode estar associada a uma maior passividade e decorrer de hábitos desenvolvidos e condicionados pelas estruturas existentes. Alguns telespectadores em primeiro lugar podem decidir ver ou não a TV e posteriormente o que ver. Logo, para promover a previsão neste contexto inicialmente a segmentação deve ser realizada com base nos padrões de exposição independentemente das escolhas realizadas. Contudo, existem outros níveis de decisão, como por exemplo, a selecção de uma tipologia, programa e fracção de programa.

A exposição poderá ser intencional ou não, assim como quando se evitam determinados conteúdos. Não existe consenso na bibliografia da especialidade, relativamente ao grau de actividade e intencionalidade subjacente às decisões do que ver. Um mesmo indivíduo pode no decorrer do tempo e em função dos contextos apresentar diferentes graus de actividade e passividade subjacente aos seus comportamentos de consumo televisivo.

Também, não existe qualquer consenso relativamente à influência das variáveis sociodemográficas na exposição televisiva. Alguns autores, salientam o papel da TV como um meio de entretenimento que abrange os mais instruídos, assim como os menos escolarizados, e outros sectores da população, ou seja, uma uniformização das motivações subjacentes à exposição televisiva. Será expectável que as principais motivações correspondam ao preenchimento de tempos de lazer, segundo nomeadamente a perspectiva estrutural, à procura de informação e numa forma de fuga à solidão.

Uma vez que a selecção em alguns contextos poderá ser intencional, mas, não, o acesso ao *media*, este processo a duas fases sublinha a debilidade de considerar as preferências e as motivações como o único componente da caracterização do consumo e nesse sentido as análises qualitativas ou menos quantitativas associadas às Ciências Sociais e à teoria de gratificações e usos do *media* são por si só insuficientes, não só para a caracterização do consumo televisivo, assim como também totalmente inadequadas para a previsão e caracterização da evolução dos índices de audiências. A exposição no global, também, pode estar associada a uma maior passividade dos telespectadores.

Apesar de alguns autores defenderem que o comportamento da audiência apresenta uma crescente complexidade a nível individual e considerando segmentos, segundo outros autores, (Cooper & Tang, 2009) actualmente, o consumo televisivo ainda se encontra fortemente condicionado pelas influências estruturais, predominando as motivações ritualísticas e assumindo os sistemas e as tecnologias de informação um papel cada vez mais relevante para a futura quantificação e caracterização da exposição televisiva.

A identificação dos métodos analíticos, mais adequados à extracção de informação dos sistemas de informação existentes apresenta consequentemente uma grande relevância. Se nos Estados Unidos se verifica esta tendência, a nível nacional com um número de canais irrisório face ao norte americano e com as divergências existentes ao nível das culturas profissionais, e não só, também se espera que o painel apresente em maior grau motivações instrumentais.

Sendo o contributo da investigação mais quantitativa determinante para a resolução de questões que preocupam a vertente empresarial deste sector de actividade, mas, não só, verifica-se, também, uma crescente necessidade pelo desenvolvimento de análises explicativas e de previsão de audiências recorrendo a abordagens Estatísticas (Weber, 2002) que permitam tirar o melhor partido dos dados que quantificam a exposição televisiva, entre os quais se insere o processo de audimetria.

Teoria de Gratificações e Usos do <i>Media</i>	Teoria Estrutural
➤ A exposição encontra-se condicionada por características individuais e de interesses por conteúdos e tipologias, estados mentais. Imperam as motivações designadas de instrumentais e a selecção ou ausência de exposição verifica-se de forma intencional. ➤ Actividade como sinónimos de utilidade, selectividade, intencionalidade e que desempenha um papel importante no processo de comunicação. ➤ Vão à fonte para avaliar as motivações, preferências, expectativas, conhecimento da grelha. Recorrem a inquéritos ou a outras abordagens mais qualitativas. Incide maioritariamente na Investigação Qualitativa.	➤ Motivações ritualísticas nos telespectadores que reflecte utilidade, mas uma menor intenção, utilidade e atenção. As necessidades estão condicionadas pelos contextos, estruturas e uma correcta estruturação da grelha viabiliza a exposição ao público desejado. ➤ Atitude passiva ou activa em ambientes estruturais, selecções podem não corresponder a reais preferências e interesses, mas decorrer de hábitos desenvolvidos e da repercussão de factores estruturais. ➤ Utilização de dados agregados para demonstrar a existência de regularidade a um nível macro de análise. Apresenta um maior poder de previsão e são úteis para explicar a exposição. Análises recorrendo à audimetria ou a outras fontes quantitativas de dados. Incide na Investigação Quantitativa.

Quadro D.2.3.1 – Teoria de gratificações e usos do *media* e Teoria estrutural da exposição televisiva
Fonte: Elaborado pela autora

Teoria de Gratificações e Usos	Teoria Estrutural
➤ Estudos Motivacionais, que se traduzem na avaliação dos motivos que levam a recorrer ao media e o que o influencia, por exemplo, o contexto social, atitudes, percepções, entre outras variáveis psicológicas e sociológicas. ➤ A influência da Psicologia nas gratificações procuradas, obtidas e nos comportamentos. Perspectiva Psicológica e os efeitos dos <i>media</i> são explicados em função dos objectivos, funções e usos do <i>media</i> controlados pelos padrões de selecção dos receptores activos. Os comportamentos são funcionais e apresentam consequências. ➤ Vertente de análise associada às Ciências Sociais, como a Sociologia e a Psicologia. ➤ Trabalhos privilegiados no Continente Europeu e na América Latina. ➤ Perspectiva de um agente activo, mas a actividade não é um conceito absoluto. Um mesmo indivíduo pode apresentar diferente grau de actividade ou passividade no decorrer do tempo e em função dos contextos. Ignoram que existem constrangimentos na exposição. Para explicar os efeitos será necessário perceber as motivações. ➤ Pode existir uma maior actividade na selecção do que na decisão de acesso ao media que decorre de hábitos desenvolvidos. Telespectadores recorrem ao media em função dos interesses, necessidades, como por exemplo, a procura de informação, redução da incerteza e a resolução de dilemas pessoais.	➤ Estudos desenvolvidos no âmbito da Gestão, Marketing, <i>Media</i> ou Métodos Quantitativos em que se avalia a repercussão de factores estruturais no consumo televisivo e nos índices de audiências. Por exemplo, a influência da sazonalidade, dia e faixa horária, disponibilidade do individuo e da grelha, variáveis sociodemográficas, acesso ao media, planeamento comercial e de estruturação da grelha, estratégias de contra-programação, tipologias programáticas, mercado de TV, notoriedade do canal, opções de media, conhecimento da programação, comportamentos de duplicação das audiências. ➤ Promovem o desenvolvimento de análises explicativas e de previsão das audiências televisivas, entre as quais se destaca a duplicação das audiências. ➤ Vertente de análise associada às Ciências Formais e à Comunicação Social. Trabalhos que incidem essencialmente nos países Anglo-Saxónicos. ➤ Agente mais passivo, existindo uma maior relação para com o media, que resulta de variáveis estruturais como a disponibilidade, e, não, para com a programação. ➤ A selecção programática, também, resulta do planeamento e conhecimento da programação. Factores estruturais influenciam a dimensão e composição da audiência, como a lealdade. Comportamento influenciado pelas estruturas. O conteúdo e a qualidade podem ter um efeito meramente marginal. Os factores estruturais influenciam fortemente a dimensão e composição da audiência, verificando-se uma exposição ou ausência de exposição não intencional e que caracteriza em maior grau a exposição global ao <i>media</i>.

Quadro D.2.3.2 – Teoria de gratificações e usos e teoria estrutural da exposição televisiva

D.3. A AUDIMETRIA E AS TEORIAS QUE EXPLICAM A EXPOSIÇÃO TELEVISIVA

A audimetria não defende, certamente, que prevalece determinada abordagem, apenas regista a exposição televisiva, apesar de alguns autores a associarem à teoria estrutural, uma vez que a perspectiva de G&U não recorre a esta tipologia de dados. Alguns autores que realizam trabalhos enquadrados na perspectiva estrutural, utilizam os dados de audimetria para confirmar as suas hipóteses de estudo, demonstrando a existência de alguma regularidade na exposição televisiva a nível agregado e a sua dependência face aos factores de planeamento. Contudo, por vezes, erroneamente concluem que a regularidade para um colectivo de indivíduos também implica a existência de regularidade a nível individual.

Segundo *Cádima (1996)*, a audimetria apresenta a limitação de não viabilizar uma análise mais qualitativa da exposição televisiva. Sendo a audimetria um sistema oneroso, inevitavelmente, foi essencialmente concebido para proporcionar ao sector os indicadores que condicionam a venda do espaço comercial e que se traduz na quantificação da exposição de forma contínua e com a maior objectividade. Para recolher meras opiniões existem abordagens alternativas e com menores custos associados.

À audimetria não interessa saber quais são os interesses ou as expectativas do telespectador. Não interessa saber, em rigor quais os seus principais gostos, quais os programas preferidos por a, b ou c. Os audímetros identificam quem vê o quê na programação. (*Cadima, 1996, p.58*)

Contudo, o autor comete uma incorrecção ao salientar que a audimetria não atribui relevância às preferências e inclusive opiniões dos telespectadores, uma vez que em determinados países europeus a tecnologia desde, há longa data, que permite a colocação de questões ao painel e, assim, recolher as suas opiniões relativamente aos conteúdos disponibilizados. Contudo, a relevância da atractividade exercida por determinados conteúdos é muito questionável quando se pretende prever a exposição televisiva, uma vez que a audimetria corresponde a um instrumento que foi desenvolvido com o objectivo de planificar e prever os indicadores que condicionam a venda do espaço comercial.

Em alguns países Europeus, como em França, os métodos de previsão de audiências quantitativos têm sido complementados com as sondagens de intenção de ver que recolhem os interesses dos telespectadores para os períodos horários que apresentam uma maior variabilidade dos dados e em geral os resultados têm sido insatisfatórios (*Zarate, 1997*).

A audimetria que corresponde a um instrumento de previsão e de planificação atribui relevância ao que é seleccionado pelos telespectadores, o que, segundo a bibliografia da especialidade, se encontra mais dependente das influências estruturais do que das motivações e interesses dos telespectadores (*Cooper & Tang, 2009*). Para *Popper e Condry (1995)*, a audimetria já permite a identificação das preferências dos telespectadores face aos programas que lhe são oferecidos, mas, não, a respeito às tendências que correspondem ao agrado dos telespectadores, o que vem contradizer, por exemplo, *Zarate (1997)*.

O que não falta são sem dúvida opiniões contraditórias a este respeito, e como *Torres (2008)* bem salienta em Sociologia imperam as divergências de opinião. Não faltam referências bibliográficas que referem a inexactidão que se observa entre o que os telespectadores manifestam como futuras selecções atendendo aos seus interesses e as selecções que posteriormente são realizadas.

Em nossa opinião, a audimetria permite identificar as selecções realizadas, que podem resultar de uma real preferência ou interesse por determinado programa ou pela escolha da alternativa que é considerada como a melhor face aos conteúdos programáticos disponíveis em determinado momento, sendo a TV consumida como um produto em si que visa preencher tempos de lazer ou como forma de evitar o isolamento e solidão, entre outros factores. A exposição poderá ser ou não intencional, assim como quando determinados conteúdos são evitados. Diversas estratégias de programação são utilizadas para tentar reter o público por um maior horizonte temporal, logo, por vezes, a exposição poderá resultar das influências estruturais no consumo televisivo, decorrente dos factores de planeamento.

Verificámos que as taxas médias de audiências apresentam uma maior estabilidade quando estão associadas a faixas horárias, e a incapacidade de manter os seus níveis de audiência quando os programas são alterados na grelha de programas. A exposição televisiva encontra-se, fortemente, condicionada pelos factores de disponibilidade e, assim, como se verifica a nível internacional devem, certamente, imperar as motivações ritualísticas.

A falta de um real interesse pela maior parte dos conteúdos transmitidos torna a exposição televisiva, fortemente, dependente dos factores de disponibilidade incutindo irregularidade no consumo televisivo a nível individual e, consequentemente dificultando fortemente as análises de previsão inclusive quando se considera um nível agregado de análise. Do ponto de vista da previsão, como consta do *Apêndice V* e da bibliografia da especialidade, as análises não devem ser realizadas com base em interesses, intenções e comportamentos passados caracterizados por frequências de exposição e baseadas em análises probabilísticas que associam o conceito de probabilidade a frequências de exposição calculadas com base num reduzido horizonte temporal.

Segundo *Cadima (1998)*, a análise dos aspectos mais qualitativos deve fazer parte de uma análise por parte dos grupos de interesse, painéis de telespectadores e conselhos de opinião. A audimetria caracteriza o comportamento de painéis de telespectadores e recorre a um método de amostragem que corresponde ao mais representativo da população, contudo os inquéritos, como inúmeros autores o demonstram, não produzem informação fidedigna. Uma vez que quanto mais as fontes de dados dependerem do factor humano mais incorrecta será a informação, têm sido desenvolvidos esforços no sentido de aperfeiçoar a tecnologia de audimetria, dando origem à denominada audimetria passiva.

Do ponto de vista da previsão determinados países, inicialmente, adoptaram metodologias híbridas que também incorporavam as sondagens de intenção de ver que correspondem a inquéritos sobre as selecções futuras. Atendendo à elevada dimensão dos erros de previsão estas abordagens acabaram por ser preteridas e foram adoptadas metodologias do âmbito do *Machine Learning* ou métodos de simulação.

Para determinados autores a audimetria já não apresenta dados incorrectos, apenas uma função que não corresponde à pretendida pela Sociologia. Por exemplo, *Sorlin (1997)* considera que os dados recolhidos pelos *media* para fotografar o seu público não são relevantes para a investigação sociológica, não por serem incorrectos, mas porque a sua função é responder a questões comerciais e não sociológicas.

A audimetria surgiu da necessidade de quantificar as audiências que corresponde a uma necessidade do sector, mas, também, viabiliza a análise de comportamentos de consumo televisivo do painel, que são os mais representativos da generalidade da população. Assim, o demonstram os inúmeros trabalhos que constam das revistas científicas internacionais, no decorrer das últimas décadas. Existem inúmeros trabalhos que acabam por concluir que a audimetria produz dados correctos, caso contrário, face à forte dependência que o sector apresenta face a estes indicadores, as empresas ligadas à medição das audiências, desde há longa data, que estariam fora do mercado. Este processo a nível nacional tem sido alvo de auditorias internas e externas e realizam-se estudos coincidentais para avaliar a correspondência de resultados. A nível internacional, por vezes, as sondagens coincidentais incidem, não só no próprio painel, como também se recorre a amostras externas ao processo.

Segundo *Cádima (1998, p.231)* a audimetria também apresenta a desvantagem de apenas registrar comportamentos: “... o telespectador estatístico é diferente do telespectador reflexivo, uma coisa é o comportamento e outra a atitude.” A audimetria nacional apenas regista comportamentos de consumo televisivo que do ponto de vista da previsão apresenta uma maior relevância, uma vez que as sondagens de intenção de ver não produzem resultados fiáveis. Recorrendo a dados de audimetria não podemos avaliar directamente determinados aspectos mais qualitativos do consumo televisivo, mas em função da evolução de determinados indicadores, certas ilações podem ser retiradas a respeito de como o público interage com a grelha de programação, complementando e servindo de suporte face a análises qualitativas posteriores em função dos objectivos das análises.

Consequentemente, os métodos analíticos, que viabilizem a extracção de informação a partir dos dados de audimetria, apresenta uma grande relevância, não só, no âmbito da previsão dos índices de audiências, como também para uma melhor compreensão de diferentes vertentes do consumo televisivo, que apresenta segundo determinados autores uma crescente complexidade.

D.4. AS ANÁLISES EXPLICATIVAS E AS ANÁLISES DE PREVISÃO DA EXPOSIÇÃO TELEVISIVA

Em inúmeras áreas as análises de previsão apresentam uma considerável imprecisão e, frequentemente, também são impossíveis se considerarmos a necessidade de calcular uma observação futura de forma “determinista”, não sendo esta a única perspectiva que pode ser considerada científica. As abordagens mais adequadas são função do grau de “maturidade” do domínio em estudo e do desenvolvimento de modelos explicativos que avaliem os factores que determinam a evolução dos dados (*Bunge, 2005*):

La predicción precisa es posible en la fisiología y la genética; otras predicciones menos exactas son posibles en la ecología, y aun mucho mas imprecisas en la teoria del aprendizaje; el minimo de precisión se encuentra en ña psicologia social. Pero en algunos campos en los que hace pocos años no eran posibles más que previsiones cualitativas se están haciendo ya predicciones quantitativas, y se tienen otras cualitativas nuevas en terrenos en los cuales, por falta de teoria, no se contaba hasta hace muy poco con ninguna posibilidad de previsión. (*p.632*)

A anterior afirmação demonstra que quanto mais madura for determinada área através da confirmação das suas teorias maior será a precisão das previsões que, realmente, sejam viáveis e mais se aprende com os erros que lhe estão associados. As ciências maduras apresentam os fundamentos das suas previsões, sendo, consequentemente, relevante uma maior compreensão da evolução dos dados. O autor, também, refere que a previsão é um conceito demasiado abrangente e que nem toda a previsão científica tem que ser, por exemplo, do tipo astronómico ou das características que viável para a análise de qualquer outro fenómeno físico.

Os modelos explicativos contribuem de diversas formas para as análises de previsão e foram considerados na presente tese. As análises explicativas contribuem para a selecção das metodologias de previsão e para o seu grau de precisão. Os modelos explicativos também podem apresentar simultaneamente capacidade de previsão se as variáveis identificadas como principais condicionantes da variável critério em análise forem conhecidas ou estimáveis com reduzidas margens de erro no momento de realizar uma previsão, para além de proporcionar as bases teóricas que fundamentam as possíveis abordagens.

Os modelos “data-mined”, ou seja, que são extraídos dos padrões observados nas bases de dados podem apresentar capacidade explicativa, mas, não, conseguir prever de forma determinista, contudo, podem viabilizar outro tipo de previsões que, também, são científicas. Segundo *Bunge (2005)*, a previsão científica não se caracteriza pela sua certeza, mas, sim, através do seu grau de adequabilidade. Considera que as previsões que são realizadas com a ajuda de regras empíricas são, frequentemente, mais exactas do que as previsões que são arduamente elaboradas com ferramentas científicas, ou seja, com leis, informações específicas e deduções. Tal é o caso com frequência das previsões meteorológicas ou das médicas e da profecia política.

Sem dúvida, que conhecer os factores que influenciam os indicadores e a exposição televisiva agregada poderá contribuir para as análises de previsão e para a maximização dos valores dos “ratings” se as estações televisivas actuarem sobre essas dimensões e se forem identificadas associações e variáveis que possam constituir variáveis de previsão. Vários autores salientam a grande relevância das análises explicativas no campo da previsão das audiências e do planeamento de *media*. Em *Patelis, Metaxiotis, Nikolopoulos e Assimakopoulos (2003)* consta que as análises exploratórias e explicativas são necessárias para a definição dos métodos quantitativos a utilizar em análises de previsão e que os principais determinantes do consumo televisivo não estão bem definidos e hierarquizados.

Segundo *Zarate (1997)*, existe a necessidade de interpretar a estrutura dos comportamentos como veículo de quantificação da exposição, uma vez que os indivíduos caracterizam-se por uma multidimensionalidade e em menor grau por uma regularidade de actos.

A propensão para realizar determinado comportamento resulta, segundo a autora, em maior grau dos diversos estímulos a que um indivíduo está sujeito em determinado momento do que por parte de possíveis “hábitos” que vão sendo adquiridos. Consequentemente, a melhor previsão poderá corresponder ao conhecimento de como se produzem e evoluem os comportamentos.

A perspectiva explicativa da exposição de gratificações e usos do *media* apresenta uma reduzida capacidade de previsão do consumo televisivo, sendo esta a perspectiva associada às sondagens de intenção de ver, onde os telespectadores manifestam as suas preferências e assumindo-se que estas coincidem com as selecções futuras. A perspectiva estrutural apresenta uma maior capacidade de previsão, desde que, também se verifique alguma regularidade na exposição televisiva. Verifica-se a necessidade de identificar as leis de estrutura que possam permanecer mais estáveis no decorrer do tempo e de desenvolvimento de modelos quantitativos, uma vez que as previsões por juízos de valor são menos precisas e existe uma maior subestimação da incerteza. Deve-se analisar e explicar as relações, os fluxos, as interdependências existentes e aceitar o conceito de probabilidade como propensão para se verificarem determinados comportamentos. O comportamento humano não é mecânico e não está sujeito a ciclos curtos, mas apresenta tendências.

No seguinte capítulo, desenvolve-se esta questão e procura-se ilustrar as divergência existentes entre as análises de previsão que podem ser realizadas nas ciências denominadas de exactas e nas ciências sociais. As considerações a apresentar fundamentam as posteriores necessidades de análise e a sequência de resultados apresentados.

APÊNDICE E _____

PERSPECTIVA DAS ANÁLISES DE PREVISÃO NAS CIÊNCIAS SOCIAIS

Neste apêndice pretende-se descrever algumas dificuldades inerentes à realização de análises de previsão nas Ciências Sociais e que permitem justificar o conjunto de análises realizadas e uma proposta metodológica que, eventualmente, poderá contribuir para que melhores resultados sejam alcançados na área dos modelos de exposição que são necessários ao planeamento de *media*.

E.1. CONSIDERAÇÕES A RESPEITO DAS ANÁLISES DE PREVISÃO NAS CIÊNCIAS SOCIAIS

As análises de previsão nas Ciências Sociais apresentam características que as diferenciam das análises que podem ser realizadas nas Ciências Exactas onde, frequentemente, podem apresentar um considerável grau de precisão e os métodos quantitativos constituem a melhor opção.

E.1.1. As análises de previsão nas ciências sociais e as suas implicações na área da audimetria

Uma das classificações mais básicas da Ciência dá-se em função do seu objecto de estudo. Deste modo, existe a divisão entre o estudo de “ideias” nas Ciências Formais e o estudo de “factos” nas Ciências Empíricas.

As Ciências Formais estão direccionadas para o desenvolvimento das ferramentas que são necessárias para se fazer ciência, como por exemplo, a Linguagem Matemática. Nesta designação podemos incluir a Estatística, a Lógica e a Ciência da Computação. Por sua vez, as Ciências Empíricas dividem-se em duas classificações: nas Ciências Naturais, que estudam a natureza, substâncias, matérias e nas Ciências Sociais, que estudam o homem e a sociedade, isto é, as suas ideias, objectivos e meios utilizados para os alcançarem, os seus comportamentos, entre outros aspectos.

As Ciências Naturais designam um ramo que estuda o universo que é entendido como sendo regulado por leis de origem natural, ou seja, os seus aspectos físicos e não os humanos. O termo Ciência Natural, também, é utilizado para distinguir as áreas que usam o método científico para estudar a “natureza”, das Ciências Sociais, que usam o método científico para estudar o comportamento e a sociedade humana; e das Ciências Formais que utilizam uma metodologia diferente.

Podemos observar certas leis na vida social, assim como na vida da natureza, contudo podem subsistir algumas dúvidas a este respeito. As ciências naturais encontram-se muito mais facilmente apoiadas na observação e na experimentação e, logo, geralmente, não implicam considerações mais rigorosas quanto à unicidade e fronteiras da ciência, sendo o método científico facilmente compatível com a metodologia específica a cada uma das subáreas que integram este grupo.

Outra classificação da Ciência envolve a sua motivação e finalidade, existindo duas classes principais: as Ciências Puras, também denominadas de Ciências Fundamentais e Básicas, que têm por objectivo o "conhecimento" em si independentemente da sua utilidade e da preocupação com as suas aplicações práticas e as Ciências Aplicadas, que estudam as formas de aplicar o conhecimento em benefício do homem. A Matemática aplicada tenta aplicar alguns modelos matemáticos teóricos à realidade, sendo isto possível apenas dentro de certas limitações e com certas restrições.

Na metade do século XX, foram desenvolvidos algumas áreas baseadas na Matemática, como por exemplo, a Investigação Operacional e a Engenharia de Sistemas. O desenvolvimento da informática deu um grande impulso a essas ciências e à teoria da computação e da informação, permitindo o estudo de sistemas complexos que ficam para além do alcance das técnicas matemáticas mais tradicionais. Na área da audimetria, certamente que a sua aplicação apresenta grandes mais valias, nomeadamente o *Data Mining* que permite a extracção de informação a partir de grandes bases de dados que surgem no meio industrial e empresarial. O desenvolvimento dessas disciplinas deixou claro que a Matemática era apenas uma das muitas Ciências Formais existentes.

Actualmente, a Matemática e a Estatística são muito aplicadas nas Ciências Naturais e nas Sociais e são essenciais para a maioria dos investigadores destas áreas, condicionando fortemente a informação que pode ser extraída das diversas fontes de dados. As Ciências Formais estudam os processos puramente lógicos e matemáticos, enquanto outras áreas, como por exemplo, a Engenharia perseguem outro tipo de objectivos. A Matemática, Estatística e a Ciência da Computação proporcionam muitas ferramentas utilizadas nas Ciências Naturais, e em outros âmbitos.

A classificação da Ciência em exactas e inexactas divide-as de acordo com o grau de precisão das descrições e previsões realizadas com base nos modelos e nas teorias científicas. As ciências exactas conseguem proporcionar resultados que apresentam algum grau de precisão a respeito dos sistemas em análise, enquanto que nas ciências inexactas as previsões são apenas direccionais e, logo, não exactas.

Em sistemas físicos podemos, por vezes, realizar análises de previsão com uma elevada precisão, como, por exemplo, na área da Astronomia e da Física. Pelo contrário na Psicologia embora se possa descrever os comportamentos e a partir deles pressupor quais serão as acções futuras mais prováveis de um indivíduo, não é na realidade possível prever-se com precisão qual será o seu próximo acto, onde e quando se irá verificar.

Conseguir prever comportamentos individuais com precisão na área da audimetria, também não é viável, uma vez que são comportamentos concretos no espaço e no tempo e o ser humano apresenta multidimensionalidade, espontaneidade, criatividade e irregularidade de actos. Os hábitos, eventualmente, desenvolvidos também não implicam obrigatoriamente uma significativa regularidade de actos no decorrer do tempo, mas, sim, uma maior propensão para que se verifiquem. Um conhecimento baseado na formulação de leis tem como pressuposto metateórico a ideia de ordem e de estabilidade do mundo, a ideia de que o passado se repete no futuro.

As Ciências Exactas onde se insere a Física, a Lógica, a Matemática e a Engenharia, entre outras, é qualquer campo da ciência capaz de expressões quantitativas, previsões precisas e métodos rigorosos de testar hipóteses, especialmente os experimentos reproduzíveis envolvendo previsões e medições quantificáveis. A Matemática, Estatística, Física, Química, assim como parte da Economia podem ser consideradas Ciências Exactas nesse sentido. O termo implica uma dicotomia entre esses campos e outros, como as ciências humanas, que possuem um carácter muito menos preciso e, logo, impera a perspectiva explicativa dos comportamentos e dos dados observados e, eventualmente, previsões meramente direccionais e probabilísticas.

Facilmente constatamos que na área das Ciências Sociais essa precisão não é viável, assim como a reprodução de experiências que implica uma realização em condições idênticas ou muito próximas, devido à multidimensionalidade que um indivíduo apresenta. De acordo com Santos (1990, p.20-21): as Ciências Sociais não têm como objectivo formular teorias explicativas; as ciências sociais não podem produzir previsões a nível comportamental, porque os seres humanos modificam o seu comportamento em função do conhecimento que sobre ele é adquirido, por se considerarem os factos sociais com natureza subjectiva; os fenómenos sociais apresentam um elevado índice de subjectividade e como tal não se deixam captar totalmente pela observação objectiva e rigorosa; as ciências sociais não são totalmente objectivas porque o cientista social, enquanto tal, não pode libertar-se dos valores que informam e condicionam a sua actividade prática.

Por último salienta que as Ciências Sociais têm como objectivo primeiro a obtenção de um conhecimento intersubjectivo, descritivo e compreensivo, em desfavor de um conhecimento objectivo, explicativo e nomotético tão próprio das ciências naturais

Termos relacionados com Ciências exactas e inexactas, mas, não equivalentes, são as Ciências Duras e as Ciências Puras. Dentro das Ciências Naturais, o termo Ciências Duras é às vezes utilizado para descrever as subáreas que dependem de dados experimentais e quantificáveis ou em que o método científico tem o seu foco na precisão e na objectividade dos resultados proporcionados. Na Física e em outros domínios das Ciências Exactas, frequentemente, encontramos sistemas deterministas que são previsíveis e, por vezes, contextos de caos determinista que, também, o poderá inviabilizar ou a previsibilidade em sistemas “indeterministas” da Mecânica Quântica e na Dinâmica. Mas, nas Ciências Sociais os comportamentos individuais são maioritariamente classificados como acausais e imprevisíveis, uma vez que variáveis não quantificáveis e previsíveis influenciam os seus comportamentos no decorrer do tempo.

Algumas problemáticas metodológicas com que, por exemplo, a Física, por vezes, se defronta apresentam algumas analogias com questões que surgem nas Ciências Sociais, contudo no âmbito da previsão existem fortes divergências. A previsão de comportamentos, inclusive no mundo animal, coloca inúmeras dificuldades, atendendo à frequente ausência de regularidade no espaço e tempo inviabilizando que resultados satisfatórios sejam alcançados. “Poderíamos opor aos fenómenos físicos, descritos pela Matemática, e ordenados no espaço e no tempo, os fenómenos estudados pelas Ciências da Vida, cuja descrição é sobretudo qualificativa”. (Serra, 2005, p.6)

Os comportamentos individuais são classificados maioritariamente como indeterminados, acausais e, também, imprevisíveis pelos científicos sociais. Esta perspectiva contrasta com a da previsibilidade dos sistemas físicos e com a perspectiva determinista ou de previsibilidade que impera nas Ciências Exactas.

Segundo *Filho (2008)*, até pelo menos ao ano de 1927, praticamente todos os físicos, quase sem excepção eram deterministas. “Podemos dizer que os cientistas continuam essencialmente deterministas, dado que as novas dimensões da Ciência, mesmo as que estão associadas a palavras como o caos e indeterminismo não impedem o domínio da natureza de que falta Bacon.” (*Serra, 2005, p.5*). Segundo *Serra (2005)*, nas Ciências Exactas tudo é determinista, e toda a gente espera que assim seja e, logo, palavras como caos e indeterminismo não servem para caracterizar o saber nas Ciências Exactas. Por outro lado, a acausalidade e imprevisibilidade dos comportamentos é algo frequentemente aceite pelos científicos sociais.

A previsão em qualquer área corresponde à temática que apresenta uma maior complexidade e em especial nas Ciências Sociais. Por conseguinte, nas análises a realizar em contextos de previsão das audiências televisivas, o objecto de estudo deverá corresponder aos índices de audiências ou considerar-se um colectivo de indivíduos, um fluxo agregado de audiências. As análises não devem incidir nos comportamentos individuais, como se verificava anteriormente quando o planeamento de media era realizado por recurso aos inquéritos.

As Ciências Naturais, que se ocupam do mundo físico, oferecem leis de grande universalidade e exactidão, apresentando um carácter determinista da realidade. Por isso, também adquire a designação de ciências duras. As Ciências Humanas, obviamente ao referir-se ao ser humano não conseguem oferecer leis tão objectivas como as anteriores, limitando-se a generalizações e a classificações da conduta, ainda que com frequência escondem as suas deficiências com a linguagem Matemática. Recebem a designação também de ciências moles, pela ideia da realidade que oferecem não ser determinista como no caso das Ciências Duras, contribuindo para uma maior dificuldade quando se pretende a realização de previsões. A Psicologia, a Economia, a Sociologia e a Antropologia são as mais representativas deste grupo.

A Sociologia é uma das áreas das ciências sociais que estuda as unidades que formam a sociedade, ou seja, estuda o comportamento humano em função do meio e dos processos que interligam os indivíduos em associações, grupos e instituições.

Enquanto que o indivíduo na sua singularidade é estudado pela Psicologia, a Sociologia apresenta uma base teórico-metodológica, que serve para estudar os fenómenos sociais, tentando explicá-los, analisando os homens nas suas relações de interdependência. Em alguns campos de estudo da Sociologia, as técnicas qualitativas, como as entrevistas dirigidas, as discussões em grupo e os métodos etnográficos permitem um melhor entendimento dos processos sociais de acordo com o objectivo explicativo.

Os Sociólogos, também, fazem uso de técnicas quantitativas de pesquisa social, como por exemplo, a inferência estatística para descrever padrões generalizados nas relações sociais. A indução também pode ser considerada uma inferência conjectural uma vez que conclui a partir da regularidade de certos factos observados, a existência de outros relacionados com os primeiros. A indução consiste em afirmar para a totalidade, aquilo que apenas foi possível observar em alguns casos e inclui a passagem do passado para o futuro que dificilmente será concretizável no campo das ciências sociais. Na inferência realiza-se uma generalização, ou seja, constroem-se proposições universais a partir de proposições particulares.

As análises de indução que podem viabilizar a previsão no campo da audimetria apenas são viáveis se existir regularidade na exposição que deverá ser avaliada com base numa amostra de indivíduos. Também ajuda a desenvolver modelos que possam contribuir para uma maior compreensão das mudanças sociais e em como os indivíduos podem responder a essas mudanças. Contudo, verificamos que no campo das audiências televisivas, a vertente quantitativa na realidade encontra-se realmente muito pouco desenvolvida a diversos níveis.

Relativamente às abordagens metodológicas aplicadas para a análise dos dados de audimetria imperam as mais tradicionais. O recurso a técnicas quantitativas e qualitativas serve normalmente, a objectivos explicativos distintos ou dependem da natureza do objecto explicado por certa pesquisa sociológica: o uso das técnicas quantitativas está associado às pesquisas macro-sociológicas; as qualitativas, às pesquisas micro-sociológicas. Entretanto, o uso de ambas as técnicas de análise de dados podem-se complementar, uma vez que os estudos micro-sociológicos podem estar associados ou ajudarem num melhor entendimento dos problemas macro-sociológicos.

Por conseguinte, a audimetria constitui uma ferramenta de análise “macro” que regista comportamentos que, posteriormente, podem ser aprofundados se necessário a nível micro pelas disciplinas e os métodos correspondentes. As análises realizadas, na presente tese, apesar de serem apenas quantitativas também evoluem de um nível mais geral de análise para um nível mais específico, recorrendo-se à análise dos índices que quantificam a duplicação de audiências para tentar explicar a evolução dos principais indicadores de audimetria e caracterizar com uma maior profundidade os comportamentos existentes e, assim, promover a previsão dos índices de audiências. Procura-se caracterizar uma informação quantitativa válida e fiável que regista a exposição do painel ao meio televisivo e que corresponde aos diversos indicadores de audimetria. Os modelos probabilísticos de previsão das audiências baseados em inquéritos não apresentam qualquer fiabilidade, assim como as sondagens de intenção de ver.

Nas Ciências Sociais é mais fácil compreender um mecanismo, ou seja, desenvolver modelos explicativos do que estabelecer uma relação de causa e efeito devido à multidimensionalidade que caracteriza um ser humano e à incidência de múltiplos estímulos nos seus comportamentos. De acordo com *Zarate (1997)*, a previsão deve ser realizada para um colectivo de indivíduos e por vezes a melhor abordagem poderá ser através do conhecimento da estrutura dos comportamentos. A evolução do ponto de vista agregado será a que condiciona a venda do espaço comercial (*Danaher, Dagger & Smith, 2011*).

Concluindo, o que está em causa neste trabalho corresponde à análise de índices de audiências que caracterizam a exposição de um colectivo e que, consequentemente, podem apresentar alguma regularidade considerando um horizonte temporal que deverá ser suficientemente abrangente. Se o comportamento colectivo, assim como o individual for caótico, dependente do acaso não será viável prever o comportamento agregado e os índices de audiências que deles resultam e que condicionam a sustentabilidade económico-financeira do sector televisivo. As análises a realizar, por conseguinte, enquadram-se nas ciências empíricas, sociais, aplicadas e inexactas e, apenas, podemos comparar diversas perspectivas de análise que não podem proporcionar uma elevada precisão ao nível dos seus resultados.

E.1.2. Considerações sobre os métodos científicos nas ciências sociais

Para *Easterby-Smith et al. (1999)* existem duas principais tradições em ciência: o Positivismo e a Fenomenologia às quais se associam determinadas suposições e implicações metodológicas. As análises de previsão das audiências televisivas enquadram-se no Positivismo, onde as propriedades do mundo social devem ser medidas através de métodos objectivos e quantificáveis e o conhecimento é alcançado pela pesquisa experimental. A realidade nesta perspectiva é estável, observável e mensurável. Pressupõe-se que os fenómenos estão sujeitos a um conjunto de leis invariáveis, tanto no mundo humano como no natural. Também rejeita que as qualidades humanas transcendem o alcance da compreensão científica.

Um método científico traduz a linha de raciocínio adoptada no processo de pesquisa. Os métodos que fornecem as bases lógicas à investigação correspondem ao dedutivo, indutivo, hipotético-dedutivo, dialético e o fenomenológico. No método indutivo o conhecimento provem da experimentação e no deductivo apenas a razão viabiliza o conhecimento. O método hipotético-dedutivo consiste na construção de hipóteses que devem ser submetidas aos mais diversos testes possíveis, para verificar quais são as hipóteses que persistem como válidas resistindo às tentativas de falseamento. Corresponde a um método de tentativas e de eliminação de erros, que não leva à certeza, uma vez que o conhecimento absolutamente certo e demonstrável não é alcançado.

As ciências empíricas caracterizam-se pela utilização do método indutivo. A procura de generalizações deve ser apoiada em evidências empíricas, pressupondo a existência de regularidades no colectivo, o que levou à necessidade de recorrer a sondagens massivas, e que no campo da audimetria se traduziu em medir os comportamentos recorrendo a painéis, isto é, a uma amostra fixa. Na indução através da observação identificam-se relações, as quais são posteriormente generalizadas. Na sua sequência, deduzem-se as suas implicações em relação a novas observações e realizam-se experiências para ver se estas se confirmam. As conclusões alcançadas por meio da indução correspondem a uma verdade não contida nas premissas consideradas, diferentemente do que ocorre com a dedução.

E.1.3. A indução e a inferência e as abordagens metodológicas

Logo, se por meio da dedução alcançamos conclusões verdadeiras, já que se baseia em premissas igualmente verdadeiras, por meio da indução chegamos a conclusões que são apenas prováveis. Comparando o método dedutivo e o indutivo, concluímos que enquanto o pensamento dedutivo leva a conclusões inquestionáveis, porém já contidas nas hipóteses, o raciocínio indutivo leva a conclusões prováveis, mas mais gerais do que o conteúdo das hipóteses. Os argumentos indutivos aumentam o conteúdo das premissas, com sacrifício da precisão, ao passo que os argumentos dedutivos sacrificam a ampliação do conteúdo para atingir a “certeza”. As metodologias de *Data Mining* enquadram-se no método indutivo e promovem a passagem de enunciados singulares, como as descrições dos resultados das observações, para enunciados universais, ou seja, hipóteses e teorias.

Usualmente denomina-se de “inductiva” uma inferência que passa de enunciados singulares a universais (*Popper, 2011*), sendo um caso especial a passagem do passado ao futuro. A indução apenas é viável se existirem regularidades no objecto de estudo, as quais se existirem são generalizadas e comprovam-se experimentalmente. Quanto maior for a dimensão da amostra sobre a qual se aceita a indução maior será a segurança com que se podem aceitar os seus resultados.

A Estatística Indutiva fornece procedimentos que orientam a realização de inferências acompanhadas de uma medida do grau de incerteza que lhe está associado. A Estatística é um ramo da Matemática que fornece os métodos que permitem obter conclusões sobre as características com interesse em conjuntos com um elevado número de elementos e associadas a fenómenos aleatórios, isto é, a situações não deterministas, em que a forma como a variável se manifesta em cada elemento é imprevisível. Como o nosso conhecimento da população é parcial e como as características com relevância para a Estatística estão associadas a fenómenos aleatórios, qualquer conclusão sobre a população terá associado um grau de incerteza.

O estudo dos fenómenos aleatórios é realizado recorrendo ao conceito de experiência aleatória. Uma experiência é qualquer processo ou conjunto de circunstâncias capaz de produzir resultados observáveis. Dizemos que uma experiência é aleatória se não pudermos prever com exactidão o seu resultado e se repetida em condições idênticas pode produzir resultados diferentes. As características fundamentais de uma experiência aleatória são, então, as seguintes:

- O aspecto imprevisível dos resultados, em que cada realização da experiência conduz a um resultado individual, mas não há conhecimento suficiente para o prever com exactidão;
- A possibilidade de repetição, ou seja, a experiência aleatória pode ser repetida um grande número de vezes nas mesmas condições ou em condições muito semelhantes;
- A regularidade estatística em que após um grande número de realizações da experiência aleatória observa-se uma forte regularidade nos resultados obtidos quando analisados em conjunto.

Facilmente se conclui que, dificilmente, se verificam algumas destas características no campo da audimetria, nomeadamente, a possibilidade de repetição da “experiência” nas mesmas condições e a regularidade dos resultados e o conhecimento *a priori* dos diversos resultados possíveis. Por exemplo, no lançamento de uma moeda para se cumprir uma lei estatística, para além, de se realizarem inúmeros lançamentos tudo deverá ser perfeito. Mas no consumo televisivo, o indivíduo e o seu entorno está submetido a todo o tipo de estímulos que modificam essa probabilidade quase que constantemente. Na moeda cada lançamento é independente, mas nas selecções programáticas cada escolha depende das anteriores. Existem condicionantes na envolvente do indivíduo que têm peso na sua área social e causam modificações na propensão ou tendência continuamente.

Nos sistemas previsíveis a informação é completamente determinada a partir de uma função geradora e das entradas da mesma, o que não se verifica na área da audimetria, uma vez que não corresponde a um sistema isolado. Contudo, inclusive em sistemas deterministas nem sempre a previsão será viável.

Os sistemas aleatórios pelo contrário tratam de incertezas, a qual pode ser classificada de duas formas: probabilística e de incerteza objectiva e mensurável ou de lógica difusa e *fuzzy* com incertezas subjectivas mais difíceis de quantificar.

Os modelos quantitativos que visam a previsão de comportamentos futuros de um conjunto de telespectadores sabendo o seu comportamento passado, podem ser diferenciados em modelos probabilísticos de inferência frequentista quando associados aos tradicionais métodos de recolha de dados, como por exemplo, aos inquéritos e em outras abordagens que foram viabilizadas pelos processos de audimetria (Zarate, 1997). A audimetria ao permitir uma análise diacrónica da audiência põe em causa a investigação tradicional e todos os modelos de planificação de media que se utilizavam anteriormente. Na *Fig. E.1.3.1* consta uma classificação das abordagens em função de serem dedutivas ou de inferência e deterministas ou probabilísticas.

	Dedutivo	Inferência
Determinista	Optimização Heurística	Query a bases de dados
Probabilístico	Decisão	Estatística
Aleatório	Simulação	Data Mining

As análises estatísticas e o *Data Mining* relacionam-se com a inferência e o método de investigação indutivo. No método dedutivo a partir do geral deduz-se o particular e na indução a partir do particular realiza-se um enunciado universal.

Quadro E.1.3.1 – Metodologias deterministas e probabilísticos

Fonte: Elaborado pela autora

A indução estatística consiste em observar as vezes que sucede determinado acontecimento e calcular a sua probabilidade e quando são viáveis será porque os acontecimentos ocorrem com uniformidade e regularidade. A indução parte do pressuposto da existência de leis universais, uniformidades, regularidades que se pensa que se irão continuar a verificar no futuro, uma vez que a evolução não é caótica, mas apenas algumas das quais se podem demonstrar.

Contudo, alguns autores defendem que não é viável passar de enunciados singulares para leis universais porque a observação desses elementos será sempre parcial e questionam a aplicação do método indutivo. A inferência indutiva corresponde ao importante ramo da Estatística que tem por objectivo fazer afirmações universais a partir de um conjunto de valores representativo sobre um universo. Para realizar este trabalho, o estatístico recolhe informações de dois tipos, experimentais em amostras e aquelas que se obtêm na literatura ou a nível empresarial. Nas análises a realizar os indicadores são medidos sobre uma amostra, um painel de telespectadores que se considera representativo do universo televisivo. As duas principais escolas de inferência correspondem à inferência frequentista ou clássica e à inferência Bayesiana.

O conceito frequentista de probabilidade envolve basicamente uma sequência de repetições para um determinado evento. A ideia da repetição justifica a denominação de “teoria frequentista”. Esta teoria baseia-se na regularidade estatística das frequências relativas e sustenta que a probabilidade de determinado acontecimento pode ser medida observando a frequência relativa do mesmo acontecimento, em uma sucessão numerosa de experiências idênticas e independentes. Para exemplificar a interpretação frequentista, considera-se uma moeda irregular, viciada lançada 1000 vezes, e observa-se a face cara 700 vezes, o que corresponde a uma probabilidade de 0,70. As análises de previsão da exposição televisiva que antecedem o desenvolvimento da audimetria baseiam-se neste conceito e, conseqüentemente não são válidas, uma vez que incidem num curto espaço de tempo, para além de outras limitações, como, por exemplo, a tentativa de identificar uma frequência que varia no decorrer do tempo.

A inferência bayesiana corresponde a um tipo de inferência estatística que descreve as incertezas sobre quantidades invisíveis de forma probabilística. As incertezas são modificadas periodicamente após observações de novos dados ou resultados. A operação que calibra a medida das incertezas é conhecida como operação bayesiana e baseia-se na fórmula de Bayes. Esta perspectiva de análise será necessária para a construção dos planos de media. A inferência Estatística associa-se com as metodologias de *Data Mining* e com a Estatística em geral.

A análise e a interpretação dos dados, associados a uma margem de incerteza, ficam a cargo da Estatística Indutiva ou Inferência, também chamada como a medida da incerteza ou métodos que se fundamentam na teoria da probabilidade. De acordo com *Hand (1998)*, a Estatística preocupa-se com a análise de dados que são recolhidos por uma razão particular e o *Data Mining* incide numa análise mais indutiva sobre dados que surgem no meio empresarial. A audimetria corresponde ao sistema de informação que viabilizou que as metodologias de *Data Mining* fossem utilizadas para descrever, prever e caracterizar os índices de audiências e os comportamentos de consumo televisivo.

Por sua vez a Lógica Difusa e a Análise Simbólica permite a representação de dados que estão associados a probabilidades difusas e a incertezas subjectivas. Perante a variabilidade do consumo televisivo, a utilização de variáveis simbólicas apresenta uma grande relevância. A literatura mostra que a estatística corresponde a um método que se aplica ao estudo dos fenómenos aleatórios e, praticamente, todos os fenómenos que ocorrem na natureza são aleatórios, como as pessoas, o divórcio, a actividade profissional, um bairro residencial, a opinião pública etc. O método Estatístico fundamenta-se na aplicação da teoria estatística da probabilidade e constitui um importante auxílio para a investigação.

Diante da impossibilidade de manter as causas constantes, os fenómenos aleatórios destacam-se porque eles se repetem e estão associados a uma variabilidade de valores. Após a ocorrência de um fenómeno aleatório, é impossível prever com precisão o resultado de nova ocorrência. Verifica-se também na repetição de um fenómeno aleatório, que os resultados se distribuem com certa regularidade, geralmente acentuada em termos de frequência. A primordial função do método estatístico corresponde à representação e explicação sistemática das observações quantitativas numéricas relativas a factores oriundos das Ciências Sociais, como o padrão cultural, comportamental, condições ambientais, físicas, psicológicas, económicas etc., que ocorrem em determinada sociedade, ou de fenómenos de diversas naturezas pertencentes a outras ciências, como na Física, Química, Biologia, entre muitas outras. São aqueles factos que envolvem uma multiplicidade de causas e por fim são representados sob uma forma analítica. Para o emprego desse método, necessariamente, o pesquisador deve ter conhecimento das noções básicas de estatística e saber como aplicá-la.

Segundo *Zarate (1997)*, quando se pretende a análise de dados nas Ciências Sociais devem-se abandonar as inferências estatísticas e a procura de relações de causa e efeito e, sim, recorrer à classificação de dados. A questão da causalidade está intimamente ligada às noções de lei e de explicação científicas e corresponde apenas a uma das formas de explicação científica. O conhecimento nas ciências sociais não se deixa apropriar segundo o modelo da causa explicativa isolada, pontual e susceptível de representação precisa através duma função matemática e apenas probabilidades podem ser calculadas. De acordo com *Almeida, Pinto e Cruzeiro (1973)*, a explicação causal traduzir-se-á antes no enunciado de uma pluralidade de condições determinantes cuja verificação simultânea confere certo grau de probabilidade à ocorrência de determinados resultados.

Alguns autores defendem que a concepção probabilística não se adequa aos acontecimentos sociais. Não devemos estimar probabilidades de uma acção por análise de alguns comportamentos isolados, mas, sim, através da interpretação dos seus comportamentos. Nas ciências exactas também encontramos referências que põe em causa as análises de previsão, como por exemplo, em *Cavalcante (2008)*, onde é referido que em sinais aleatórios é teoricamente impossível prever o valor exacto do sinal em determinado instante a partir de uma observação em um instante diferente ou apenas conhecendo o valor da abcissa. Em resumo um sinal aleatório não tem uma formulação matemática como uma função. Segundo este autor, os modelos matemáticos com que usualmente trabalhamos não são capazes de traduzir o comportamento dos sistemas e sinais reais que encontramos na prática, por isso recorre-se a análises probabilísticas para componentes que sejam apenas um reflexo do que ocorre no mundo real. Esta constatação também se aplica ao campo da audimetria, onde lidamos com comportamentos humanos que, segundo inúmeros autores, são acausais e imprevisíveis. Contudo, as únicas previsões possíveis também não são as que alguns autores classificam de “deterministas”, ou seja, que pretendem o cálculo de uma grandeza quantitativa num momento concreto. Para que seja possível prever os comportamentos devem-se observar regularidades que também se manifestem em hábitos e que podem ser identificadas recorrendo a análises de duplicação das audiências. Contudo, nem sempre o hábito se associa a uma forte regularidade dos resultados observados.

O comportamento colectivo poderá não ser caótico, mas, sim, o individual. Se existirem regularidades, só podem ser identificadas quando se recorre a um grande número de pessoas e a um longo horizonte temporal, que não é o mais habitual para um investigador, ou seja, no curto prazo podem não se manifestar e existe a necessidade de recorrer a painéis.

E.1.4. A abrangência do conceito de “previsão”

Frequentemente, associamos ao conceito de previsão apenas a possibilidade de antever o valor de uma observação concreta no espaço e no tempo e em especial nas Ciências Exactas. Contudo, as previsões nem sempre podem e devem ser precisas. Podemos identificar outras perspectivas de análise que também são científicas e que contribuem, positivamente, para os diversos domínios em estudo.

A projecção estocástica será a menos precisa e traduz-se numa enumeração de possibilidades, como, por exemplo, a passagem de um estado x para um estado y em sucessivas etapas ou intervalos de tempo. Uma árvore de probabilidades corresponde à única previsão possível em muitas situações da vida quotidiana e em especial nas Ciências Sociais. Apesar de ser qualitativa corresponde a uma previsão científica se as possíveis transições se apresentarem em função de leis que sejam simultaneamente enunciadas.

Nas Ciências Sociais as várias probabilidades estimadas podem diferir significativamente e, logo, podem estar justificadas nas áreas em que será relevante descartar as situações menos prováveis de ocorrer. Em inúmeros contextos interessa essencialmente inferir a respeito do desenvolvimento mais provável dos acontecimentos na próxima unidade temporal. Neste contexto estará em causa prever uma tendência na alteração de um sistema, a sua direcção mais provável e não os seus estados futuros com uma grande precisão. De acordo com *Bunge (1981)*, as Ciências Sociais beneficiam, certamente, com um estudo teórico e empírico mais aprofundado das probabilidades de transição, permitindo assim, prever a direcção da evolução mais provável dos acontecimentos e que resultados mais precisos possam ser alcançados.

Isto é o que se faz em física atómica e na vida quotidiana, com a particularidade que neste último caso a projecção estocástica é geralmente qualitativa e não quantitativa e sem aprofundar ou debilmente fundamentada e sem serem baseadas em teorias científicas. A projecção estocástica é na realidade cientificamente respeitável e devemos ter isto presente se realmente se pretender uma maior aceleração da evolução das ciências não físicas e não exactas.

Independentemente do nível de agregação dos dados considerado, as análises de previsão podem ser baseadas em leis estocásticas, ou não, e que será função do conhecimento disponível e do objectivo. Segundo *Bunge (1981)*, a projecção estocástica, ou seja, estatística para o colectivo ou probabilística a nível individual não deve ser considerada como menos importante e interessante do que a projecção não estocástica ou determinista. As análises devem ser realizadas com o contributo das teorias e sobre a base de generalizações estatísticas que são viáveis a nível empírico. A previsibilidade não suporta nem implica a retroactividade e a projectabilidade é relativa ao conhecimento que está disponível e não intrínseca aos acontecimentos em análise.

Consequentemente, as análises explicativas são essenciais para as análises de previsão e em função dos contextos podemos adoptar as perspectivas de previsão que lhe são realmente mais adequadas. Segundo a bibliografia disponível nas Ciências Sociais devem ser privilegiadas as análises colectivas e sobretudo as qualitativas. Contudo, no campo da audimetria é necessária uma quantificação da exposição com a maior precisão possível e segundo se demonstra as análises mais qualitativas não constituem uma boa opção e não disponibilizam a informação necessária à construção dos planos. Será necessário avaliar diversas perspectivas quantitativas que também são função dos contextos em análise.

E.1.5. Algumas dificuldades que a previsão coloca nas ciências sociais

As análises de previsão nas Ciências Sociais levantam inúmeras dificuldades e que se caracterizam nos seguintes parágrafos, assim como as respectivas implicações e as necessidades de análise que originam.

➤ A inferência indutiva poderá viabilizar a previsão de observações futuras a partir dos dados históricos se existirem regularidades.

Uma inferência indutiva viabiliza a passagem de enunciados singulares para enunciados universais e corresponde à lógica da probabilidade.

Segundo *Reinchenbach (1930)* este método determina a verdade das teorias científicas e para *Popper (2011)* apesar de não serem estritamente válidas podem atingir determinado grau de certeza ou de probabilidade, ou seja, correspondem a inferências prováveis, mas os enunciados científicos também apenas podem alcançar contínuos graus de probabilidade. A indução implica a necessidade de se verificarem regularidades nos acontecimentos observados, que se podem eventualmente manifestar em hábitos de consumo televisivo.

O conceito de hábito, nem sempre, implica necessariamente que os comportamentos se verifiquem com uma forte regularidade em determinado contexto espacial e temporal, apenas que existe uma maior propensão para poderem ocorrer. Segundo *Wartofsky (1973)*, constituir um hábito transcende a realização de uma acção de forma parecida no decorrer do tempo, ou seja, é um conceito mais geral do que os seus casos ou circunstâncias de ocorrência e vai além de uma descrição sintética dos casos em que determinado facto se consegue observar. Apresentar o hábito de realizar algo torna-o provável, mas poderá não se concretizar em determinado contexto que lhe seja equivalente, ou seja, podem-se criar circunstâncias que impedem a sua concretização.

Logo, apesar de alguns segmentos poderem apresentar determinados hábitos de consumo televisivo nem sempre se vão concretizar e podem reduzir a precisão das previsões consideravelmente. Quando as acções resultam do acaso não podemos enunciar qualquer regra, mas se lhe associarmos hábitos denota um carácter regulado sobre as quais podemos eventualmente representar determinadas relações invariantes através de um enunciado simbólico universal. Um conhecimento indutivo será apenas provável. Para a inferência inductiva devem existir leis, uniformidades na exposição do público face a determinado suporte. As regras de inferência lógica não abrangem os casos de forma necessária, mas apenas probabilística.

Consequentemente, será necessário realizar análises à duplicação das audiências e ver em que contextos poderá existir uma maior regularidade na exposição aos suportes televisivos e avaliar a influência das variáveis estruturais a este nível e que constam das bases de audimetria. Deve-se avaliar a predisposição para existir uma maior regularidade na exposição ao media e aos suportes.

➤ Em análises de previsão deve ser considerada uma elevada dimensão da amostra e que deverá corresponder a um colectivo de indivíduos que seja representativo da população em estudo e não a observações individuais, uma vez que os comportamentos individuais são classificados como imprevisíveis e acausais. Os comportamentos individuais apresentam características que podem tornar a sua previsão por métodos quantitativos inapropriada, como por exemplo, circunstâncias não previsíveis do dia a dia de cada pessoa. Alguns autores salientam que estas situações podem não ter uma grande repercussão nos resultados, mas que devem ser avaliadas (*Salzinger, 2005*):

The so called “broken-leg” phenomenon is invoked when you are trying to predict whether a particular man is going to see a particular movie. All the information that you have about his predilections may still lead you to the wrong prediction because he has a broken leg and therefore simply cannot go to movies at all. (You also have to assume that there is no wheelchair around and that movie houses do not provide special places for wheelchairs, as they do nowadays.) psychology does not have many “broken legs” ... there are few conditions that trump universal generalizations: Nevertheless, such special conditions ought to be investigated to determine how often they do occur. (*p.93*)

Apenas recorrendo a análises de duplicação das audiências será viável avaliar o grau de regularidade da exposição e na *questão II* em avaliação, na presente tese, referem-se os diversos contextos de exposição que afastam as selecções realizadas dos conteúdos que podem ser considerados como mais apelativos pelos telespectadores. Os comportamentos sociais também são classificados, frequentemente, e como acausais e imprevisíveis a nível individual, mas da sua agregação podem resultar comportamentos não caóticos que manifestem a existência de determinadas regularidades. As projecções devem ser realizadas para um colectivo, que poderá constituir um sistema, mesmo que não estejam fortemente inter-relacionadas as partes que o integram.

Segundo *Bunge (1981)*, um comportamento regular pode resultar de comportamentos aproximadamente independentes. Todo o componente individual poderá parecer indeterminado, mas na realidade corresponder a uma estrutura geral. Toda a regularidade estatística representada por uma tendência ou uma lei refere-se a uma estrutura colectiva; mas nem toda a regularidade referente a um colectivo é estatística, e nem mesmo probabilística. Denominamos estatística a uma projecção que se baseia em regularidades estatísticas. Uma projecção estatística é colectiva no sentido de se referir a um agregado estatístico, ou seja, a uma colecção de entidades e apresentam alguma causalidade interna, por muito pouco causais que possam ser em outros aspectos.

A explicação das regularidades pode requerer uma análise dos componentes num nível diferente, mas a previsão estatística não terá tal em consideração e referir-se-á ao todo e no seu próprio nível. A explicação das regularidades apenas será possível recorrendo a um nível de análise mais desagregado e inferior ao nível das análises de previsão. As análises de duplicação das audiências realizadas também decorrem desta necessidade. As previsões a nível individual são estocásticas, sendo necessário probabilizar os seus comportamentos. Numa experiência aleatória também não conseguimos prever o resultado de um ou de um número reduzido de acontecimentos que são finitos e conhecidos a prior, mas sim apenas de uma sucessão na qual se irá manifestar a lei probabilística que o rege.

De acordo com *Popper (2011)* o resultado de um lançamento de um dado é aleatório e não é viável prever o seu resultado em cada lançamento. Segundo *Bunge (1981)*, as previsões não se referem a observações precisas e individuais, mas, sim, a propriedades colectivas e a características de longo prazo e as probabilidades podem ser apenas, por vezes quantificadas como base nas frequências observadas.

Encontramos el azar en sentido objectivo, puede decirse, cuando nuestras estimaciones probabilitarias resultan confirmadas, del mismo modo que encontramos regularidades causales cuando las predicciones que deducimos de leyes vienen a corroborarse. (*Popper, 2011, p.192*)

A identificação das regularidades estatísticas no colectivo pode ser alcançada recorrendo a metodologias de *Data Mining*, sendo necessário avaliar o seu potencial de aplicação recorrendo a dados agregados para o universo ou por segmentos, uma vez que o painel não é homogéneo quanto aos perfis de consumo televisivo. As regularidades convertidas em probabilidades permitem que sejam realizadas as inferências. As previsões têm que ser realizadas com base em indicadores de audimetria ou em fluxos de audiências e também se encontram certamente dependentes da qualidade dos dados e das tecnologias disponíveis em determinado momento.

As melhores previsões podem ser alcançadas com melhores dados e métodos de processamento dos mesmos, sendo os sistemas de informação e as metodologias estatísticas de análise de dados determinantes. Um fenómeno social apresenta em geral propriedades colectivas e a partir do estudo de características individuais não chegamos ao colectivo. Constatações que se verificam a nível agregado não podem ser associadas a nível individual.

Por exemplo, ao analisar dados agregados referentes a países, Estados é difícil determinar se as mesmas relações existem em níveis desagregados, ou seja, em indivíduos e lares. O mesmo sucede no campo da audimetria, o que pode viabilizar a possível previsão de indicadores agregados, ainda que a exposição ao nível individual seja caótica e não previsível. Ao contrário do que *Fernandes (2001)* defende a estabilidade que por vezes se observa nos valores dos ratings em períodos adjacentes com reduzidos valores de duplicação não corresponde a uma inconsistência, mas sim advém deste factor.

Segundo *Dilão (1995)*, as propriedades macroscópicas podem não estar relacionadas com as propriedades individuais:

Quando se junta um grande número de sistemas, as propriedades macroscópicas ou colectivas do sistema composto não estão, em geral, relacionadas com as propriedades dos seus constituintes individuais. Neste caso, o sistema composto é um sistema complexo.

A evolução dos índices de audiências, também, não está relacionada com propriedades individuais. O que é relevante corresponde à estrutura geral do comportamento da audiência e não a um comportamento de um indivíduo numa experiência concreta. As situações individuais não são relevantes. Uma vez que não se pode induzir correlações colectivas, a partir de correlações individuais, também, é relevante a identificação dos contextos em que os indicadores agregados traduzem os comportamentos que se verificam a um nível mais desagregado. Consequentemente, nas análises de duplicação deve-se avaliar em que medida as regularidades observadas são ou não consistentes com a evolução dos principais indicadores de audimetria.

➤ A amostra deve abranger um longo horizonte temporal, logo é necessário recorrer a painéis e à audimetria e a métodos de amostragem probabilísticos, uma vez que correspondem aos únicos que permitem realizar inferências precisas.

Nas Ciências Sociais não podemos estimar a probabilidade de uma acção analisando apenas uns quantos comportamentos isolados, como se verifica nas análises probabilísticas de previsão das audiências que recorrem ao conceito de probabilidade enquanto frequência. Com uma aprofundada análise do painel de audimetria poderá ser muito difícil detectar regularidades de comportamentos entre indivíduos e a grelha de programas. Com os métodos de medição das audiências televisivas que não viabilizam uma análise diacrónica da audiência será difícil detectar uma frequência de contacto, uma vez que apenas se realiza uma contagem do número de vezes que se fez uma selecção no decorrer das últimas semanas, ou seja, num curto espaço de tempo. Torna-se difícil captar a ordem ou regularidades estatísticas nos comportamentos observados em curtos espaços de tempo, logo torna-se indispensável adoptar a audimetria.

➤ O colectivo pode apresentar comportamentos diferenciados, mas as características sociodemográficas poder ser insuficientes. Será necessária a identificação das melhores abordagens de segmentação do painel e a avaliação da influência sociodemográfica na exposição ao media e na selecção televisiva, assim como na evolução dos principais indicadores de audimetria.

➤ Nas análises de previsão apenas podemos considerar um subconjunto de variáveis, as que são mais relevantes, conhecidas e privilegiando a informação que consta das bases de audimetria. Contudo, variáveis não quantificáveis, ambíguas e não controláveis influenciam os resultados e a inferência nunca poderá ser precisa e apresentar a objectividade e rigor de uma experiência física.

Segundo *Meehl (1954)*, o ser humano apresenta uma grande dificuldade em combinar informação de forma correcta e fiável, e, logo, as análises estatísticas são essenciais para a avaliação da relevância das diversas variáveis explicativas e de previsão.

Será necessário identificar as variáveis de audimetria que exercem uma maior influência na evolução dos principais indicadores e avaliar se são conhecidas ou previsíveis quando se pretende realizar uma análise de previsão. Nas análises de regressão linear as variáveis independentes são deterministas e no campo da audimetria, existem poucas variáveis conhecidas ou previsíveis com alguma fiabilidade no momento de realizar uma previsão. Nas análises multivariadas realizadas pretendeu-se verificar quais são as variáveis conhecidas que são mais relevantes do ponto de vista explicativo e/ou de previsão das audiências. A indução consequentemente nunca será exacta e não tem o rigor de uma experiência física, sendo necessário quantificar a ordem de grandeza do erro em função das abordagens implementadas. Os telespectadores apresentam liberdade e neles incidem inúmeras variáveis que influenciam permanentemente as suas acções.

➤ Os telespectadores modificam o comportamento no decorrer do tempo. Verificamos que existe interdependência entre sucessos e que a exposição a determinados suportes pode condicionar exposições sucessivas.

Consequentemente, é necessário avaliar os comportamentos de duplicação, entre os quais os efeitos de herança e a lealdade ao canal, tipológica ou programática. Quando se constroem os planos de media é necessário que as abordagens de previsão tenham em atenção o carácter dinâmico da evolução dos dados e que consequentemente o conhecimento anteriormente adquirido seja capitalizável de forma máxima para horizontes temporais posteriores. “Factors that help to make predictions change over time, requiring that formulas be modified to accommodate those changes.” (*Salzinger, 2005, p.93*). Relativamente à continuidade e descontinuidade dos comportamentos existem diferentes perspectivas a respeito dos factores que apresentam capacidade explicativa e de previsão. Segundo *Foxall (2009)*:

The prediction of such behaviours follows easily enough from consideration of the contingencies alone (at least for aggregates, not necessarily for individuals), but an explanation of the change itself requires consideration of the process of comparison and recognition that must precede the change. ... Selective perceptions, beliefs and desires must be used as part of the explanation of such behaviour. (p.218) ... we must be clear that beliefs and desires are not causal: they cannot enter into an experimental analysis. Our only reason for using them must be that the observed behaviour can be fully explained only by ascribing them. (p.219)

➤ A envolvente de um telespectador influencia a sua exposição. O ser humano apresenta uma multidimensionalidade e encontra-se condicionado pelo entorno social que modifica o seu comportamento (*Salzinger, 2005*):

The behaviour of those whose behaviour we try to predict might well be a function not of the variables we invoke but a consequence of others people's behaviours, others histories of those people whose behaviour we are trying to predict, etc.” ((p.95). “Behaviour is a function of both contingent and happenstance reinforcement. (p.96)

O indivíduo apresenta uma multidimensionalidade e em determinado comportamento participam inumeráveis causas. A vontade humana é livre e as suas manifestações numa acção concreta não são totalmente previsíveis, assim como não existem regularidades invariáveis nos fenómenos sociais.

Una fuente importante de fallas en la predicción es el conjunto de suposiciones acerca de la naturaleza del objeto cuyo comportamiento ha de predecirse. Por ejemplo, puede ocurrir que creamos que el sistema en cuestión está suficientemente aislado de las perturbaciones exteriores, cuando en rigor estas cuentan a la larga; dado que la aislación es una condición necesaria de la descripción del sistema con ayuda de un puñado de enunciados de leyes, no debiera sorprender que sea tan difícil predecir el comportamiento de sistemas abiertos tales como el océano, la atmósfera, el ser vivo, o el hombre. Puesto que la predicción científica se funda en las leyes científicas, hay tantas clases de predicciones como clases de enunciados nomológicos. Algunas leyes nos permiten predecir resultados individuales, aunque no sin error si la predicción se refiere al valor de una cantidad. Otras leyes, incapaces de decirnos nada acerca del comportamiento de los individuos son en cambio la base para la predicción de algunas tendencias globales y propiedades colectivas de colecciones numerosas de elementos similares: son leyes estadísticas. Las leyes de la historia son de este tipo; y por esto es casi imposible la predicción de los sucesos individuales en el campo de la historia, pudiendo preverse solamente el curso general de los acontecimientos. (*Bunge, 2005, p.32*)

Os comportamentos futuros resultam de uma associação entre a regularidade e a multidimensionalidade que caracteriza o ser humano. Um dos problemas é que não se cumprem as regularidades estatísticas, pelo menos no curto prazo. Como não se pode prever acaba-se por interpretar o comportamento e analisam-se as possíveis causas pelas quais em determinado momento um indivíduo apresenta a propensão para realizar determinado comportamento.

➤ Cada indivíduo está indeterminado e é imprevisível, mas valores médios ou agregados podem ser mais facilmente previsíveis. Contudo, equivalentes valores médios podem reflectir uma grande diversidade de estados de um sistema, e, logo, não reflecte uma descrição objectiva do comportamento do objecto em análise.

Consequentemente, recorre-se à Análise Simbólica que permite uma melhor representação da evolução dos dados. Deve-se avaliar a evolução de variáveis simbólicas que são mais representativas da evolução de determinado indicador, em vez de considerar apenas valores médios ou variâncias dos valores originais.

➤ Uma vez as previsões exactas não são viáveis e em geral as abordagens conduzem a margens de erro que podem ser significativas, é necessário promover, também, as análises explicativas e identificar as abordagens analíticas que o viabilizam.

Nenhum sistema de previsão será exacto porque estamos a lidar com comportamentos sociais. Será necessário ter em atenção em análises de previsão a sua rapidez, se apresenta conhecimento que seja capitalizável para momentos posteriores, baixo custo, prever apenas ou também interpretar e ter em atenção as alterações de comportamentos e da evolução dos dados, a sazonalidade e as abordagens que considerem a envolvente do indivíduo.

➤ Nas análises de previsão não podemos prever a exposição considerando a interpretação frequentista da probabilidade. Será necessário avaliar que definições da probabilidade se adequam aos comportamentos sociais. Os comportamentos resultam das vivências diárias e nem sempre de actos semelhantes praticados no passado.

Logo, são inúmeras as limitações das análises de previsão na área das Ciências Sociais. Cada indivíduo está indeterminado, a sua conduta não é previsível e para alguns autores apenas faz sentido em falar de propensão ou tendências. Os indivíduos podem modificar o seu comportamento e atitudes face aos programas e estes últimos na realidade não são independentes. Poderá existir regularidade estatística nos comportamentos sociais quando tomados no seu conjunto, mas apenas se poderá captar com um grande número de pessoas e considerando grandes horizontes temporais.

Certamente, que existem fenómenos sociais onde ainda que se verifique essa regularidade não se poderá captar, uma vez que será necessário considerar um horizonte temporal muito abrangente o que não será viável por motivos económicos, entre outros. Uma experiência nas ciências sociais não tem o rigor da física, nem a inferência poderá ser tão exacta.

APÊNDICE F

APRESENTAÇÃO DE RESULTADOS

Neste apêndice constam os resultados das análises realizadas, a nível explicativo e de previsão dos “*ratings*”, onde se privilegia a análise dos indicadores de audimetria que condicionam maioritariamente a actividade de planeamento de *media*, nomeadamente a análise dos indicadores de duplicação das audiências entre os suportes publicitários e a análise das taxas médias de audiências programáticas, comerciais e para o total da TV. Em função destes resultados no *Capítulo VI* é proposta uma sequência de análises a utilizar em determinadas etapas do planeamento de *media* e no *Capítulo V* constam os principais resultados e a sua discussão.

As perspectivas teóricas e empíricas explicativas da exposição televisiva ainda apresentam significativas limitações a diversos níveis, proporcionando no geral uma visão não integrada e insuficiente dos diversos factores que determinam o consumo televisivo. A frequente mencionada crescente complexidade ao nível da caracterização da interacção dos telespectadores com o meio televisivo, também, acarreta uma maior dificuldade quanto ao futuro desenvolvimento de modelos explicativos e de previsão dos “*ratings*” ajustados a diversos níveis de agregação dos dados ou dos fluxos de audiências entre os suportes publicitários. Quanto maior for o nível de desagregação dos dados considerado, mais difícil ou inclusive inviável será a previsão da exposição televisiva.

No âmbito da previsão quantitativa das audiências televisivas, que é essencial para o planeamento de *media*, os reduzidos trabalhos publicados na literatura internacional da especialidade incidem maioritariamente em abordagens metodológicas tradicionais e apresentam significativas limitações, desconhecendo-se os factores que condicionam a evolução das audiências e que possam simultaneamente apresentar capacidade predictiva.

Pretende-se o desenvolvimento de análises explicativas e em determinados contextos de previsão dos principais indicadores de audimetria, nomeadamente das taxas médias de audiências para o total da TV, *ratings* programáticos e de períodos horários e comerciais, *shares* de audiências e para a variável *Spot_Ratio%*. Simultaneamente, inter-relaciona-se a evolução de alguns destes indicadores com a evolução dos fluxos de audiências entre os suportes publicitários e em especial com os indicadores acumulados de audimetria, que permitem avaliar a existência e caracterizar os comportamentos estruturais de consumo televisivo. Os *ratings* condicionam a venda do espaço comercial, contudo a estrutura de preços, também, é definida por análise de determinados indicadores que permitem uma interpretação da sua evolução. As análises de duplicação das audiências são essenciais para a construção das campanhas publicitárias e no seguinte *Quadro F.1* constam os seus principais objectivos.

Objectivos das Análises de Duplicação Das Audiências
A avaliação da Regularidade da Exposição sem a qual a inferência indutiva não é viável
Avaliação da Existência dos Comportamentos Estruturais no painel; factores que os condicionam; De que forma interagem e a sua influência na evolução dos <i>ratings</i> , permitindo uma avaliação qualitativa da sua evolução e a construção de um modelo explicativo do consumo televisivo
A análise da duplicação das audiências contribui para a selecção dos suportes a incluir em planos de <i>media</i> e para a avaliação da viabilidade Das análises de previsão de valores de <i>ratings</i> e/ou da exposição televisiva e na Identificação das melhores perspectivas de análise no âmbito da construção dos planos de <i>media</i>

Verificamos que são escassos os trabalhos que visam avaliar o contributo de áreas metodológicas mais recentes e adequadas à análise das bases de dados de audimetria, como por exemplo, o *Data Mining* ou a Análise Simbólica, considerando quer as análises explicativas, como também, as análises de previsão. Apenas foi possível identificar um reduzido número de aplicações de metodologias de *Data Mining* proposicional.

Quadro F.1 – Objectivos das análises de duplicação das audiências
Fonte: Elaborado pela autora

Do ponto de vista explicativo, a utilização das bases de dados de audimetria apresenta grandes vantagens a diversos níveis, e considerando as análises de previsão da exposição televisiva são essenciais. Nas bases de audimetria consta uma grande diversidade de dados que apresentam uma significativa variabilidade e nas quais as variáveis, frequentemente, se inter-relacionam de forma não linear. Logo, o *Data Mining* e a análise Simbólica correspondem às áreas metodológicas mais adequadas, uma vez que permitem a representação dos dados em estruturas mais ricas face às tradicionais e a avaliação das interações existentes entre as diversas variáveis estruturais com influência na evolução dos indicadores.

Não foi possível identificar nas revistas internacionais da especialidade trabalhos que visem a análise das bases de dados de audimetria nacionais. Apenas foram identificados contributos muito pontuais realizados em departamentos de Engenharia ou Matemática. A inviabilidade de partir de determinados pressupostos ajustados à realidade do sector televisivo Português, contribuiu para a necessidade de serem realizadas um vasto conjunto de análises. Nesse sentido, foi necessário realizar todas as análises exploratórias de caracterização da evolução dos principais indicadores e contextos que condicionam a actividade de planeamento de *media* e a comercialização dos espaços comerciais. Esta tese, face aos recursos disponíveis procura identificar, sugerir e explorar algumas das potencialidades de áreas metodológicas mais recentes que possam contribuir para a avaliação da repercussão de variáveis estruturais em diversos índices de audiências e a correspondente previsão.

Ainda que as audiências comerciais sejam o indicador mais relevante, a sua dependência face à envolvente programática requer que, também, sejam realizadas análises neste âmbito. As agências de planificação de meios e os restantes profissionais deste sector em geral apenas dispõem de dados agregados de audimetria para realizar as suas análises, sendo esta a tipologia de dados que também foi utilizada para a realização desta tese e que é a adequada às análises de previsão ou explicativas dos índices de audiências. Menciona-se no *capítulo VI* os contextos em poderá apresentar um maior interesse recorrer aos dados dos ficheiros de base de audimetria.

Nos *Apêndices F.1 a F.7* constam os resultados das análises univariadas e multivariadas realizadas aos diversos índices de audiências programáticos e comerciais com base em dados agregados de audimetria para o universo e por segmentos sociodemográficos e recorrendo a metodologias, quer do âmbito da previsão, quer da segmentação, ou seja, classificação e agrupamento de dados.

Os estudos de audiências, como consta da bibliografia da especialidade, têm como objectivo quantificar a exposição televisiva, que por sua vez se encontra fortemente dependente dos métodos de medição das audiências e de análise dos dados. A mais valia do trabalho realizado associa-se com a definição e ilustração de métodos analíticos mais recentes, conciliados com abordagens mais tradicionais que permitam uma análise mais adequada dos dados de audimetria, contribuindo para a clarificação de questões relacionadas com o consumo televisivo e com a previsão dos índices de audiências. Pretendemos verificar se através de metodologias mais recentes é viável maximizar as percentagens de variância explicada dos principais indicadores e promover a identificação dos factores com uma maior repercussão, assim como as respectivas interações a nível comercial, programático e na interface de ambas as vertentes.

No âmbito da duplicação das audiências, também, se recorre a dados mais adequados à sua análise e inexistentes ou minoritariamente explorados na literatura internacional da especialidade, por exemplo, relativamente aos efeitos de herança e aos índices de lealdade programática, assim como do ponto de vista analítico se privilegiam os métodos não lineares. A análise de determinados comportamentos estruturais é relevante para as análises de previsão de séries de TV ou de qualquer outro programa que apresente uma sequência de episódios. Em determinados contextos pode ser necessário avaliar quais são as componentes da audiência, a sua dinâmica e se são previsíveis.

F.1. ANÁLISES EXPLORATÓRIAS AOS INDICADORES DE AUDIMETRIA QUE CONDICIONAM A CONSTRUÇÃO DOS PLANOS DE MEDIA

F.1.1. Análise de taxas médias de audiências para o total da TV

A caracterização das taxas médias de audiências para o total da TV é determinante para o desenvolvimento de modelos explicativos e de previsão das audiências televisivas. As taxas médias de audiências que condicionam a comercialização da publicidade televisiva dependem do respectivo *share* de audiência face à dimensão da exposição global. Consequentemente, iniciam-se as análises realizando uma análise o mais exaustiva possível a esta variável, mas condicionadas pelas características dos dados que foram disponibilizados. Pretende-se caracterizar a sua evolução, nomeadamente a sua variabilidade, promovendo, simultaneamente, a identificação dos factores estruturais com uma maior repercussão a ambos os níveis. A bibliografia internacional demonstra que persistem inúmeras incógnitas relativamente aos factores que determinam a exposição ao meio televisivo (*Cooper & Tang, 2009*).

No *Apêndice F.1.1.1* consta uma síntese do estado da arte e caracterizam-se alguns dos factores que condicionam a sua evolução, fazendo, simultaneamente um levantamento das necessidades de análise e estabelecendo-se as inter-relações existentes com os capítulos posteriores. Segue-se a definição dos objectivos das análises a realizar (*Apêndice F.1.1.2*) e a descrição das bases de dados e das metodologias efectivamente utilizadas (*Apêndice F.1.1.3*). No *Apêndice F.1.1.4* constam os resultados repartidos por quatro subcapítulos: análises que visam a caracterização da variável *TVR_Prog%* quanto à sua evolução e factores que a condicionam (*Item I*), ou seja, a análise da variável quando associada à duração de um programa; análises de caracterização da evolução de *TVR_Período%*, ou seja, quando a variável se encontra associada a faixas horárias (*Item II*); a análise sociodemográfica em que se avalia para cada segmento a dimensão e a variabilidade da sua exposição ao meio televisivo (*Item III*) e a influência dos espaços comerciais na sua evolução (*Item IV*).

F.1.1.1. Introdução

Este capítulo incide na análise do indicador de audimetria que traduz a exposição do painel ao meio televisivo, sob várias vertentes. Os autores *Cooper e Tang (2009)* referem que alguns factores de cariz estrutural, demográfico e de motivações explica no máximo 30,3% da variabilidade inerente à exposição ao meio televisivo no mercado norte-americano com base num inquérito realizado a alunos e funcionário da universidade. Os autores salientam que as análises que pretendem avaliar com um maior rigor os factores que condicionam a exposição televisiva, devem ser mais quantitativas e incidir em outras tecnologias de medição das audiências, como os diários de exposição televisiva. Porém, foi demonstrado no *Apêndice A* que esta fonte de dados apresenta maiores limitações e uma menor validade face às bases de dados de audimetria. O desenvolvimento dos actuais sistemas de informação irá certamente permitir que no futuro seja viável explicar com um maior rigor o comportamento dos telespectadores a diversos níveis. No entanto, nesta tese não se pretende explicar as motivações e interesses dos telespectadores, mas, sim, a análise dos indicadores de audimetria.

Actualmente, permanecem inúmeras incógnitas quantos aos factores que condicionam a selecção deste *media*. Os factores que influenciam a exposição televisiva podem ser subdivididos em estruturais ou individuais. Os dados de audimetria em especial ao nível dos dados dos ficheiros de base podem viabilizar a avaliação das interacções existentes entre estas duas tipologias de determinantes do consumo televisivo. A exposição televisiva por sua vez condiciona a eficácia das diferentes estratégias de planeamento e a evolução dos *ratings* programáticos e comerciais.

As publicações identificadas nas principais revistas internacionais da especialidade que visam a análise deste indicador de audimetria são reduzidas, assim como os trabalhos que caracterizam e explicam os padrões de exposição ao meio televisivo, independentemente das selecções programáticas realizadas pelos telespectadores. Os autores *Barwise e Ehrenberg (1988)* referem que a dimensão da audiência total apresenta uma evolução regular, sendo determinante a sua influência nos padrões de consumo.

Webster e Litchy (1991) sublinham a importância dos padrões de disponibilidade dos telespectadores quando se pretende prever os valores de *ratings*. Salientam que esta variável poderá explicar a disparidade existente entre as preferências e as selecções efectivamente realizadas. Referem a forte sazonalidade inerente ao consumo televisivo e a diferenciada evolução desta variável no decorrer do dia e da semana.

Também, tem sido constatado que as faixas horárias que não estão associadas ao horário nobre apresentam um menor número de telespectadores, mas uma maior regularidade e consistência nos padrões de disponibilidade e de exposição dos telespectadores. *Webster e Wang (1992)* confirmam esta tendência e salientam que os factores estruturais, em particular a disponibilidade do público são as variáveis explicativas que condicionam em maior grau os padrões de consumo televisivo.

Nas últimas décadas, por vezes, tem sido constatado que esta variável não depende do conteúdo programático transmitido, apresentando valores muito semelhantes ano após ano para idênticos períodos horários. Assim sendo, defende-se que a audiência adopta hábitos de consumo, apresentando um padrão regular de exposição. Por oposição, outros autores defendem que o conteúdo programático e em especial o conhecimento que o público detém da grelha de programas influencia esta variável. Por exemplo, *Eastman e Newton (1998)* analisaram a influência das promoções de programas nos valores dos índices de audiências. Concluíram que, por vezes, os telespectadores organizam-se para poder assistir a determinados programas. No entanto, não podemos generalizar estas tendências à globalidade do painel e faixas horárias. Actualmente, existindo uma maior oferta ao nível das opções de *media* e de canais televisivos será questionável a validade das conclusões de trabalhos realizados durante a década de 80 e 90 e que constam da bibliografia internacional.

Em Portugal, no período que decorre entre 1999 e 2004, assistimos a um acréscimo de 8% nos *shares* de audiências dos canais por TV Cabo e no final de 2003, 27% dos lares portugueses eram assinantes da TV Cabo. Os comportamentos de consumo televisivo certamente não permanecem alheios à evolução verificada ao nível dos meios de comunicação e de enquadramento económico, laboral, cultural e social.

Com dados mais recentes e que resultam de inquéritos, *Cooper e Tang (2009)* apenas conseguem alcançar uma capacidade explicativa de 30% da evolução dos dados nos USA. Para os autores, os resultados ficam muito aquém do esperado e salientam a necessidade de novas abordagens serem utilizadas por forma a promover uma melhor percepção de como a audiência interage com este *media* e compreender a evolução da exposição global.

O mais lógico será em princípio pressupor que, actualmente, o consumo televisivo resulta de ambos os factores, ou seja, de uma certa passividade e hábitos adquiridos por parte de alguns segmentos de telespectadores, assim como em determinados contextos, especialmente quando associados a programas que não são frequentemente transmitidos, de um comportamento mais activo do público. Em princípio determinados programas, como os jogos de futebol, as galas, algum filme em estreia ou um concurso de Eurovisão podem justificar que a audiência altere as suas rotinas de consumo televisivo. A dimensão desta repercussão, também, poderá em princípio ser função das faixas horárias consideradas e dos canais, uma vez que estes diferem quanto à composição sociodemográfica e psicográfica do seu público. *Nikolopoulos et al. (2006)* realizaram um trabalho que teve como objectivo caracterizar a repercussão nos valores dos *shares* programáticos gregos que resulta da transmissão de alguns eventos desportivos.

Provavelmente, determinados períodos horários podem apresentar uma menor tendência para que se verifique um notório acréscimo nos índices de audiências como consequência da introdução de programas de maior popularidade, mas, sem dúvida, que os canais também reservam estes programas para os períodos de maiores audiências. Os programas que normalmente constam da grelha de programas e que têm como objectivo promover a fidelização da audiência devem em princípio apresentar uma menor repercussão na variabilidade desta variável, sendo esta essencialmente função dos padrões de disponibilidade e de selecção do *media* e em menor grau dos conteúdos programáticos transmitidos.

Interessa portanto verificar, com base em dados nacionais, o grau de estabilidade na evolução desta variável e em que medida alguns programas podem influenciar a variabilidade observada. Posteriormente, nos capítulos seguintes, também ao nível da evolução dos *ratings* e dos *shares* de audiências.

Certamente que podem surgir diferentes cenários na evolução das variáveis nestes contextos, como, por exemplo, os acréscimos verificados na dimensão das taxas médias de audiências para o total da TV podem-se repercutir de forma diferenciada nos restantes indicadores em função da capacidade dos programas para conseguir retirar audiências aos canais concorrentes.

Consequentemente, no *Apêndice F.4*, analisam-se os fluxos de audiências entre os canais para tentar avaliar o grau de independência entre o público de cada um dos canais e a sua repercussão na evolução dos *ratings*. No *Apêndice F.5* constam algumas análises às exposições sociodemográficas com o intuito de caracterizar o perfil de audiências de alguns programas, tipologias e canais e avaliar o seu grau de “lealdade”. Talvez, as variáveis sociodemográficas influenciem a predisposição para se desenvolverem comportamentos de lealdade a determinados programas e tipologias programáticas e, consequentemente, também, a determinados canais. Quanto maior for a fidelização da audiência à programação e ao canal, menor será a repercussão da contra-programação e mais facilmente se poderá prever os índices de audiências em determinados contextos e períodos horários.

As características sociodemográficas podem influenciar o grau de passividade, não só, relativamente à selecção programática, como, também, quanto à possível existência de rotinas no que diz respeito à frequência e períodos horários seleccionados. Talvez, as maiores oscilações nos valores das taxas médias de audiências para o total da TV se possam atribuir a um subconjunto de segmentos de telespectadores. Caracterizando o público alvo de determinado canal e identificados os segmentos que apresentam uma maior ou menor propensão para desenvolverem rotinas de consumo e exposição ao *media* podemos, posteriormente, avaliar o grau de variabilidade desta variável decorrente da introdução de programas que podem apresentar uma maior popularidade.

Quando as análises incidirem nos factores que condicionam a exposição televisiva, podemos recorrer aos dados dos ficheiros base de audimetria e realizar uma análise de segmentação do painel em função dos seus padrões de disponibilidade, exposição e tendencia para desenvolverem rotinas de consumo televisivo.

Com base nas análises sugeridas verificar-se-á se existem segmentos diferenciáveis e identificáveis e se as características sociodemográficas realmente constituem, por si só, factores explicativos do comportamento da audiência a este nível. Os telespectadores podem-se diferenciar quanto às motivações subjacentes ao consumo televisivo, o qual pode constituir um veículo de informação, um meio de preenchimento de tempo livre ou vazio ou para a satisfação de outras necessidades.

Uma vez que para a realização desta tese dispomos de dados agregados de audiências televisivas podemos comparar e avaliar a evolução desta variável para os segmentos representados nas bases de dados de audimetria, identificando quais são as características sociodemográficas que contribuem em maior grau para a variabilidade associada aos indicadores que caracterizam o universo televisivo.

Os padrões de disponibilidade e de exposição, também, estão condicionados, por exemplo, pela climatologia, alternativas lúdicas existentes ao nível dos *media* e, não só, lealdade tipológica, programática e ao canal, ciclo laboral e escolar, agregado familiar, entre muitos outros factores, nem sempre quantificáveis. Consequentemente, a previsão dos valores desta variável com reduzidas margens de erro será aparentemente cada vez mais difícil de concretizar e dependente dos períodos horários, dias e meses ou dos segmentos considerados.

Provavelmente, determinadas faixas horárias, programas ou conjunto de telespectadores podem apresentar uma maior estabilidade na evolução deste indicador. A relevância desta variável para o desenvolvimento de modelos explicativos de *ratings* justifica a sua análise, sendo este o objectivo deste capítulo. No seguintes parágrafos caracterizam-se alguns dos factores que segundo a literatura da especialidade exercem uma maior repercussão na exposição televisiva. Os factores que influenciam a evolução desta variável e o seu grau de variabilidade são inúmeros dificultando a sua previsão.

Alguns dos factores que condicionam a exposição televisiva podem ser agregados em dois subconjuntos: os factores estruturais e os factores individuais.

Nos seguintes parágrafos apresentam-se algumas considerações referentes aos itens apresentados, definindo-se as necessidades de análise e as inter-relações existentes com os apêndices posteriores.

□ **Factores Estruturais**

I. A Disponibilidade da Audiência

Cooper e Tang (2009) referem que a investigação realizada neste âmbito tende a operacionalizar a audiência disponível como a percentagem da audiência que selecciona este *media* e não do número de pessoas que apresentam disponibilidade para o realizar.

Partem do pressuposto de que se o público estiver disponível, também, assistirá à TV, o que poderá não se verificar necessariamente. Actualmente, existem outras opções ao nível dos *media* ou de outras actividades lúdicas que podem ser seleccionadas em detrimento do televisor, especialmente à medida que tais opções constituam alternativas funcionais ao meio televisivo.

A Internet tem sido apresentada como um potencial concorrente da TV, assim como a rádio e os vídeos jogos, mas *Cooper e Tang* acabam por concluir que se verifica o inverso. Por exemplo, a Internet é uma variável positivamente correlacionada com a exposição televisiva. Consequentemente, a escolha da TV e da Internet funciona de modo complementar, satisfazendo necessidades funcionais diferentes.

As motivações do público, também, podem influenciar a regularidade observada na sua evolução. Espera-se a existência de alguma regularidade nos horizontes temporais em que determinados segmentos de telespectadores apresentam uma maior disponibilidade para assistir à televisão, e consequentemente, também, ao nível da correspondente exposição televisiva. O grau de regularidade da exposição ao televisor deverá depender das características sociodemográficas dos telespectadores, entre outras, assim como dos períodos horários, notoriedade do canal, popularidade dos programas e da estruturação da grelha. Nas análises a realizar aos índices de “repeat viewing” pretende-se verificar qual é a percentagem do painel que apresenta uma exposição mais regular em função das faixas horárias e tipologias programáticas transmitidas.

Neste apêndice, pretendemos comparar os segmentos em função da regularidade da sua exposição; avaliando-se o grau de variabilidade inter-anual e, também, para o período em análise, assim como promover a identificação dos factores com uma maior influência a este nível. Os segmentos que apresentem uma evolução mais variável talvez devam ser objecto de análises posteriores, que permitam avaliar a probabilidade de selecção deste *media* perante determinados contextos, desde que correspondam aos segmentos alvo dos canais em análise e o seu valor comercial o justifique. As motivações para assistir à TV em horizontes temporais não habituais depende da disponibilidade para ver TV.

Consequentemente, a transmissão de programas de grande popularidade que podem levar à maximização das taxas médias de audiências para o total da TV deverá apresentar uma maior variabilidade em função dos factores de disponibilidade do painel que em geral não recorre a este *media* no decorrer dos períodos horários equivalentes. Estas situações devem exigir a utilização de diferenciados instrumentos analíticos face aos contextos em que se verifique uma selecção mais regular. A capacidade para atrair um maior número de telespectadores, também, deverá depender dos conteúdos programáticos e do canal de transmissão.

Os padrões de disponibilidade do público influenciam os comportamentos de duplicação de audiências, ou seja, os efeitos de herança, a lealdade programática, tipológica e ao canal que também são objecto de análise nesta tese (*Apêndice F.3 a F.5*).

II. Os Novos *Media*

Cooper e Tang (2009) salientam que as audiências não devem ser conceptualizadas como sendo passivas ou activas, mas, sim, eventualmente como “activas em âmbitos estruturais”. O público será simultaneamente activo e passivo, uma vez que os *media* convergentes suprem uma série de necessidades dentro e através de processo de comunicação. Os autores defendem que a dicotomia de activo e passivo deveria ser, em princípio, abandonada reconhecendo a significativa influência da estrutura nas escolhas realizadas. As estruturas ao nível dos *media* podem modificar-se no decorrer do tempo, mas as estruturas de disponibilidade e de acesso permanecerão no decorrer de horizontes temporais delimitados, rendimento e padrões de actividade diária. Os indivíduos procuram determinadas estruturas para conseguir lidar com um cada vez maior número de opções ao nível dos *media* e dos conteúdos.

Para o período em análise, a maior parte do painel apenas tem acesso aos canais generalistas. A estrutura não deverá ser encarada como uma característica mais passiva de utilização dos *media*, mas, sim, como uma das influências que inegavelmente existe neste domínio. O painel de audimetria nacional apresenta em menor percentagem lares com acesso à TV Cabo, levando a pressupor que o grau de penetração da Internet, também, será diminuto, assim como a real concorrência exercida ao nível dos restantes *media*, que em geral apresentam diferentes vertentes funcionais, segundo a literatura da especialidade.

Os *shares* de audiência dos canais por TV Cabo, permanecem muito aquém dos valores observados para os canais generalistas, mantendo estes últimos a liderança quanto aos suportes mais pretendidos do ponto de vista comercial. A TV representa essencialmente um meio de entretenimento e de informação em horizontes temporais que não incentivam à realização de actividades alternativas, sejam estas de lazer ou laborais, e para alguns segmentos representa um veículo de preenchimento de tempo de lazer.

Considerando a concorrência ao nível do orçamento e sendo a televisão generalista um serviço gratuito, as alternativas lúdicas concorrentes a considerar são função da época do ano, inculcando uma considerável sazonalidade ao consumo televisivo.

A nível global a pressão exercida por concorrência indirecta ou potencial poderá ser menos significativa para grande parte do painel, estando a exposição televisiva condicionada em maior grau pela disponibilidade do público. Pretendemos verificar se a exposição se encontra condicionada em maior grau por factores de disponibilidade, acabando por incutir alguma regularidade neste indicador em determinados horizontes temporais. A análise de fluxos de audiências e de outros índices de duplicação das audiências poderá confirmar a existência de alguma actividade nas selecções realizadas, e em geral talvez o reduzido envolvimento e fidelização das audiências para com este *media*. Será necessário avaliar se a repercussão das decisões de ligar e desligar o televisor na evolução dos *ratings* é superior face à que resulta da transição dos telespectadores entre os canais.

III. O Canal

As divergências existentes quanto ao posicionamento, segmentos de mercado-alvo e quotas de mercado de cada estação televisiva, obviamente, influênciam a evolução dos principais iniciadores e a variabilidade decorrente das alterações realizadas ao nível da estruturação das suas grelhas e decorrentes da imagem transmitida ao público. Neste sentido a par da repercussão da variável canal em *TVR%*, no *Apêndice F.1.2* realizaram-se algumas análises que pretendem avaliar como o desempenho dos canais influencia a evolução das taxas médias de audiências dos programas e no *Apêndice F.6* se condiciona a relevância das variáveis estruturais no âmbito programático e no *Apêndice F.7* no âmbito das audiências comerciais. No *Apêndice B* foi caracterizado o sector televisivo nacional e referem-se as diferenças que existem ao nível das estratégias de programação de cada um dos canais, as respectivas estratégias de contra-programação, a evolução dos seus *shares* e as perspectivas futuras para este sector.

IV. O Acesso

Na literatura da especialidade, essencialmente referente à década de 80 e 90, defende-se que a existência de um maior número de canais, em geral, não origina um maior consumo televisivo global, nem uma maior diversificação no conjunto de programas seleccionados.

Os telespectadores podem se identificar apenas com um subconjunto mais restrito de programas para facilitar o processo de selecção. A nível nacional, verificamos que apesar de uma parte do painel ser subscritor da TV Cabo, os índices de audiências são mais reduzidos para estes canais. A repercussão da programação dos canais por cabo, nos valores dos *ratings* dos canais generalistas e ao nível dos efeitos de herança, será analisado no *Apêndice. F.4*. No entanto, espera-se que a sua repercussão não seja, para o período em análise, muito significativa.

V. O Conhecimento da Programação

A influência das promoções de programas, também, deverá ser avaliada, especialmente quando são transmitidos programas que geram observações atípicas. Os dados de audimetria nacionais, inclusive ao nível dos ficheiros base de audimetria, pela sua estruturação dificultam a realização de análises a este nível. Uma inserção caracterizada como sendo uma promoção na base de dados Publitext não se encontra identificada pelo programa a que corresponde, sendo necessário assistir aos vídeos que lhe estão associados. Aparentemente, a sua repercussão será mais relevante face à transmissão de programas mais populares. Com base na análise dos valores de *ratings* dos eventos desportivos podemos verificar se os acréscimos verificados em *TVR%* decorrentes da sua transmissão são superior nos canais privados ou no canal estatal, uma vez que a audiência dos canais privados apresenta uma maior e mais regular exposição televisiva, estando mais facilmente exposta às promoções programáticas.

□ Factores Individuais

I. As Motivações dos Telespectadores

As motivações da audiência, na literatura da especialidade, são diferenciadas em motivações instrumentais e ritualísticas (*Cooper & Tang, 2009*). As motivações instrumentais reflectem a existência de actividade e a procura de informação ou de outros conteúdos, enquanto que as motivações ritualísticas traduzem uma selecção mais passiva deste *media* com o intuito de preencher os tempos livres e vazios.

Regra geral, as motivações ritualísticas devem, certamente, imperar com excepção da selecção de alguns conteúdos programáticos, como, por exemplo, os espaços informativos e determinados programas de desporto. Nas análises a realizar pretendemos verificar se à audiência do canal estatal podemos associar em maior grau as motivações instrumentais, através da análise da evolução de TVR% quando se transmitem determinados eventos desportivos, e se os espaços informativos podem transportar por inércia as audiências para os programas sucessores e durante que horizonte temporal.

II. As Características Sociodemográficas

Na bibliografia da especialidade não existe consenso relativamente à influência das características sociodemográficas nos padrões de exposição televisiva, e posteriormente quanto à forma de interagir com o *media*. Em princípio espera-se que as classes sociais mais baixas, as maiores faixas etárias e o sexo feminino assistam em maior grau à televisão. As classes sociais mais altas devem apresentar uma maior predisposição e recursos para contratar canais por cabo apesar da sua menor disponibilidade, verificando-se o inverso nos mais jovens e nas classes sociais mais baixas. Consequentemente, em função do peso de cada variável de segmentação no público-alvo de cada canal, tal condicionará a variabilidade observada em TVR% e ao nível dos *ratings* face à introdução de programas que o painel possa considerar mais apelativos e as análises a realizar vão nesse sentido.

Segundo *Cooper e Tang (2009)*, as características sociodemográficas associam-se em menor grau com o nível de exposição televisiva, apresentando uma maior correlação com as motivações e a disponibilidade. Salientam que estas variáveis não apresentam uma grande repercussão na exposição televisiva, uma vez que nas suas análises apenas a faixa etária apresentou alguma influência. As análises a realizar pretendem avaliar qual será a repercussão destas variáveis nos indicadores de audimetria. Ao nível da variável TVR%, verificamos que os diversos segmentos diferenciam-se significativamente quanto à variabilidade e evolução observada neste indicador. A avaliação da repercussão destas variáveis nos valores dos *ratings* consta do *Apêndice F.1.2*.

Também, foi possível associar algumas tipologias, programas e canais aos segmentos que constam das bases de audimetria, recorrendo a metodologias simbólicas e de *Data Mining*. No *Apêndice F.5* verificou-se que as variáveis sociodemográficas em parte condicionam as associações existentes. Contudo, as características sociodemográficas não são, certamente, o único factor a considerar quando se pretende discriminar os padrões de consumo televisivo. Será interessante verificar, se a nível nacional, se verificam as mesmas tendências que constam da bibliografia internacional.

F.1.1.2. Objectivos

Os factores que influenciam a exposição televisiva são inúmeros, apresentando interdependências que são função do horizonte temporal em análise. Alguns dos principais factores que podem ser analisados recorrendo a dados agregados de audimetria correspondem ao factor tempo e a características de planeamento e dos conteúdos da grelha. Os horizontes temporais condicionam o grau de disponibilidade do público e a predisposição para se desenvolverem rotinas de consumo televisivo, inculindo uma maior regularidade em TVR%.

Consequentemente, neste apêndice pretendemos realizar a análise da exposição ao meio televisivo em função do tempo, através da caracterização da evolução de TVR% sob várias vertentes. Pretendemos verificar qual é o grau de estabilidade da exposição total ao *media* e o seu grau de independência face aos conteúdos transmitidos. A relevância desta variável, que certamente deverá constituir um factor a ser incluído em modelos explicativos e de previsão das audiências televisivas, justifica a sua caracterização.

Apenas foram disponibilizados os dados agregados de audimetria e que corresponde à tipologia de dados a que têm acesso as agências de meios. Uma vez que esta tese não se restringe apenas à avaliação deste factor, sendo necessário a caracterização das audiências sob várias vertentes com vista ao desenvolvimento de modelos explicativos e de previsão das audiências, vão ser privilegiadas as seguintes análises que *a priori* devem ser consideradas como sendo as mais relevantes:

(1) Caracterizar a evolução da taxa média de audiência para o total da televisão quando está associada a programas, uma vez que apresenta uma considerável associação linear com o valor das taxas médias de audiências programáticas, e, logo, constitui uma perspectiva de análise que talvez possa contribuir para a previsão dos valores *ratings* com um maior grau de precisão.

(2) Cálculo e caracterização da evolução das taxas médias de audiências para o total da TV considerando os valores para faixas horárias e verificando em que medida a sua dimensão é independente do conteúdo programático transmitido.

Se os valores observados não diferirem significativamente para os vários anos em análise podemos estar perante um painel que apresenta em maior grau um comportamento mais passivo e segundo hábitos desenvolvidos no decorrer do tempo. Contudo, o conceito de hábito não tem que ser necessariamente equivalente a uma forte estabilidade nos padrões de consumo televisivo. Determinados períodos horários podem-se diferenciar a este nível, assim como o contributo dos diversos canais para a variabilidade observada.

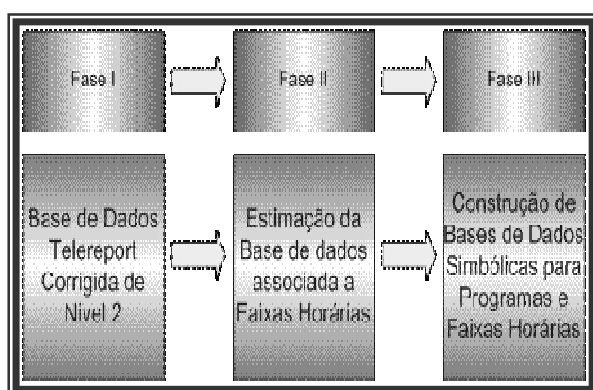
(3) Avaliação da repercussão de determinados programas na sua evolução, nomeadamente de galas, eventos desportivos, entre outros.

A influência de determinados programas poderá estar condicionada por parte dos períodos horários, canal de transmissão ou contra-programação. Também, é relevante avaliar como alterações na evolução desta variável se repercutem nos valores dos *ratings* e *shares* programáticos, não só do canal, como também dos canais concorrentes e este aspecto será considerado no *Apêndice F.1.2*.

(4) Diferenciar os segmentos representados nas bases de dados, atendendo à evolução e variabilidade da sua exposição ao *media* representada através da variável *TVR_Período%*.

Talvez as maiores oscilações existentes na exposição televisiva para o universo televisivo correspondam à repercussão de um reduzido número de segmentos. Pretende-se segmentar as bases de dados em função da dimensão da exposição, evolução e grau de estabilidade. Ainda que uma melhor percepção e caracterização do processo de exposição ao meio televisivo seja relevante para o sector audiovisual, do ponto de vista organizacional, também, é relevante a previsão deste indicador considerando o universo televisivo e por segmentos atendendo à sua repercussão na evolução dos *ratings*. Nesse sentido, apesar dos múltiplos factores que condicionam a exposição televisiva, as análises a realizar pretendem avaliar o grau de envolvimento da audiência com o *media*, favorecendo ou não uma maior regularidade neste indicador e que está obviamente dependente dos canais, factores temporais e talvez das características sociodemográficas dos telespectadores.

F.1.1.3. Bases de dados



Os dados em análise correspondem às taxas médias de audiências para o total da TV associadas aos programas (variável designada de *TVR_Prog%*) ou alternativamente a períodos horários com uma duração de 30 minutos (variável designada de *TVR_Período%*).

Fig. F.1.1.3.1 – Estruturação das bases de dados associadas à variável TVR%
Fonte: Elaborado pela autora

Para a análise dos valores de *TVR%* programáticos recorreu-se ao nível 1 e ao nível 2 de um programa no *software* Telereport. As bases de dados de nível 2 tiveram que ser corrigidas, agregando-se os programas que apenas apresentam uma única parte programática. Na Fig. F.1.1.3.1 consta a descrição do processo de construção das bases de dados que viabilizaram a análise deste indicador.

A variável *TVR_Prog%* que consta do *software* Telereport designa as taxas médias de audiências para o total da TV quando está associada ao universo televisivo ou a segmentos na área programática. Representa o número de telespectadores que contactaram com a TV durante a transmissão de determinado programa ou evento:

$$Tvr\_Evento\% = \frac{\text{Total de segundos contactados no Total da televisão}}{\text{Duração do evento} \times \text{N.º de indivíduos do alvo}}$$

Esta variável será classificada em global ou parcial e designada por *TVR_Prog%* (*P*) ou *TVR_Prog%* (*T*) em função do nível programático considerado, quer dizer, considerando a totalidade da duração de programa ou apenas cada uma das partes que o integram. A variável *TVR_Prog%* poderá referir-se ao universo televisivo quando consideramos a globalidade do painel ou apenas a cada um dos segmentos que integram as bases de dados disponibilizadas, sendo as variáveis de segmentação as seguintes: o sexo, as faixas etárias, a situação no lar, as regiões geográficas e as classes sociais (ver *Apêndice I* para descrição das bases de dados).

A variável *TVR_Período%* corresponde ao indicador calculado para intervalos horários de 30 minutos e para a totalidade da emissão diária (7H-25H). Calcula-se através da seguinte expressão:

$$Tvr\_Periodo\% = \frac{\text{Total de segundos contactados no Total da televisão}}{30 \text{ minutos} \times \text{N.º de indivíduos do alvo} - \text{Universo ou por segmento considerado}}$$

Nas bases de dados simbólicas que foram construídas, os conceitos ou objectos de segunda ordem no âmbito programático correspondem a programas associados a determinado dia do período laboral ou de fim-de-semana e segundo a correspondente faixa horária quando o mesmo programa é transmitido, por exemplo, antes e após o espaço informativo noite, como, por exemplo, Espaço_informativo_noite_Segunda_20H ou Espaço_informativo_noite_Fim-de-semana_21H.

Os conceitos ou objectos de segunda ordem no âmbito temporal correspondem a períodos horários de 30 minutos de duração, como, por exemplo, 19H-19.30H_Segunda ou 19H-19.30H_Fim-de-semana, diferenciando-se os valores de fim-de-semana dos valores dos dias laborais ou em função dos dias.

Nesta tese pretende-se a análise do indicador que traduz a percentagem da audiência que efectivamente selecciona em determinado instante este *media*, uma vez que condiciona a evolução dos restantes índices de audiência e a relevância das interações existentes entre as restantes variáveis explicativas e os indicadores de audimetria.

Considerou-se na apresentação dos resultados, em maior grau, o período que decorre desde 2000 a 2004, pelas razões enumeradas anteriormente (*Apêndice B*) e considerando desde 15 de Março a 15 de Maio para cada ano que integra a amostra desde 2000 a 2008.

Apresenta uma maior relevância verificar em que medida a instabilidade no cenário televisivo pode influenciar a variabilidade deste indicador do que recorrer a um horizonte temporal que apresente uma maior consolidação do desempenho das estações televisivas no cenário nacional dos *media*.

Atendendo à sazonalidade do consumo televisivo, o horizonte temporal seleccionado em cada ano, é considerado como sendo característico do consumo televisivo a nível médio. Pretendeu-se minimizar a repercussão exercida por determinadas alternativas concorrentes ao meio televisivo, tentando aproximar em maior grau os padrões de disponibilidade do público com os reais padrões de selecção do meio televisivo. Os profissionais do meio consideram estas semanas como sendo representativas do consumo televisivo médio e influenciadas em menor grau pela sazonalidade existente nesta área.

F.1.1.4. Metodologias

I. Análises de Caracterização de TVR_Prog% e TVR_Período% (Universo)

Para prever este indicador e seleccionar as metodologias mais apropriadas será necessário em primeiro lugar caracterizar a sua evolução e grau de variabilidade, recorrendo a análises gráficas, a algoritmos de *clustering* de séries temporais de *Data Mining* Temporal e a algoritmos de *clustering* simbólicos.

Para analisar o grau de semelhança existente entre as séries temporais originais recorreremos nomeadamente à distância eucladiana e ao coeficiente de correlação linear de pearson, aplicando-se, posteriormente, o algoritmo K Média ou o *Clustering* Hierárquico. A distância eucladiana entre duas séries será representada por

$$D(Q,C) = \sqrt{\sum_{i=1}^N (q_i - c_i)^2}$$

e o coeficiente correlação linear de pearson representado por

$$r_{ij} = \frac{\sum_{t=1}^L p_i(t) p_j(t)}{\sqrt{\sum_{t=1}^L p_i(t)^2 \sum_{t=1}^L p_j(t)^2}}, 1 \leq j \leq N, 1 \leq i \leq M$$

onde, $P_j(t) = \sum_{k=1}^{N_j} p_{jk}(t)$, $1 \leq t \leq L$, $1 \leq j \leq N$

e $p_i(t) = \frac{(p_i^*(t) - m_i)}{\sigma_i}$, $1 \leq t \leq L$, $1 \leq i \leq M$

sendo $P_j(t)$ a série temporal média do *cluster*, r_{ij} o coeficiente de correlação linear de pearson entre uma série temporal e a série média e $p_i(t)$ a série estandardizada.

O número de *clusters* avalia-se através da análise em Componentes Principais.

Para as séries temporais simbólicas as semelhanças são avaliadas recorrendo a inúmeros indicadores, entre os quais a distância de Hausdorff, L1 e L2. Considerando n objectos descritos por p variáveis da tipologia intervalo $Y_j: E \rightarrow B_j: x_i \rightarrow Y_j(x_i) = x_{ij} = [\alpha_{ij}, \beta_{ij}]$ definem-se os p índices de dissemelhanças $\delta_1, \dots, \delta_p$ no conjunto β_j entre $x_{uj} = [\alpha_{uj}, \beta_{uj}]$ e $x_{vj} = [\alpha_{vj}, \beta_{vj}]$ com base na distância Hausdorff representada por

$$\delta_j(x_{uj}, x_{vj}) = \max\{|\alpha_{uj} - \alpha_{vj}|, |\beta_{uj} - \beta_{vj}|\}$$

a distância L1 representada por

$$\delta_j(x_{uj}, x_{vj}) = |\alpha_{uj} - \alpha_{vj}| + |\beta_{uj} - \beta_{vj}|$$

a distância L2 representada por

$$\delta_j(x_{uj}, x_{vj}) = (\alpha_{uj} - \alpha_{vj})^2 + (\beta_{uj} - \beta_{vj})^2$$

Posteriormente, combinam-se os p índices $\delta_1, \dots, \delta_p$ para obter uma medida global de dissemelhança em E .

$$d : ExE \rightarrow R^+ : (x_u, x_v) \rightarrow$$

$$d(x_u, x_v) = \left(\sum_{j=1}^p \delta_j^2 (x_{uj}, x_{vj}) \right)^{1/2}$$

II. Análise dos valores de *TVR_Período%* associados a Segmentos Sociodemográficos

Os valores de *TVR_Prog%* ou *TVR_Período%* para o universo televisivo resultam da interacção existente entre os correspondentes valores para os vários segmentos sociodemográficos.

Para identificar quais são os segmentos que apresentam uma evolução da variável em estudo mais semelhante à do universo televisivo através do *indexing* de séries temporais ou que mais se aproximam entre si através do *clustering* de séries temporais, podemos recorrer ao DM Temporal ou a metodologias de análise Simbólicas.

Provavelmente a variabilidade inerente a *TVR_Período%* (Universo) poderá ser atribuída a um subconjunto de segmentos. Na área do *Data Mining* Temporal promove-se o desenvolvimento de metodologias que comparam a evolução de séries temporais no domínio temporal, identificando as mais semelhantes.

Como algoritmos de *clustering* de ST podemos recorrer ao algoritmo K Média ou ao *clustering* Hierárquico, entre outros. Alternativamente pode-se realizar a sua passagem para domínios não temporais, ou seja, de frequência ou tempo-frequência que permitam a aplicação de algoritmos proposicionais de *clustering*.

Por exemplo, recorrendo às transformadas de Fourier ou a *Wavelets*. Uma vez que as séries temporais de *TVR_Prog%* e de *TVR_Período%* associadas aos diferentes segmentos apresentam igual comprimento, unidade de medida, linhas base e distribuição das observações não se justifica em princípio a passagem das séries temporais para domínios não temporais ou recorrer a metodologias de alinhamento de séries, por exemplo, ao “Dynamic Time Wrapping”.

A semelhança entre as séries temporais pode ser avaliada no domínio temporal por utilização da distância euclidiana ou calculando os coeficientes de correlação e posteriormente aplicando metodologias hierárquicas de *clustering* ou algoritmos de *Data Mining*.

Alternativamente, podemos também recorrer aos métodos de cálculo de (dis)semelhança e *clustering* simbólicos, correspondendo a base de dados ao intervalo de variação da variável *TVR_Período%* para determinados períodos horários.

Do ponto de vista simbólico existem inúmeras medidas de dissemelhança entre objectos simbólicos booleanos e probabilísticos.

Nesta base de dados, uma vez que as variáveis são da tipologia intervalo e adoptando como objectos de segunda ordem as faixas horárias com 30 minutos de duração estamos perante objectos simbólicos booleanos.

F.1.1.5. Resultados

Para avaliar a viabilidade de previsão deste indicador será necessário, em primeiro lugar, caracterizar a sua evolução, verificando quais são os factores responsáveis pelas maiores amplitudes de valores. Analisa-se esta variável, atendendo ao seu valor para as durações programáticas parciais e globais, e posteriormente para os períodos horários. Pretende-se a análise da sua evolução em função dos seus valores no contexto programático, uma vez que condiciona a previsão das taxas médias de audiências programáticas, que por sua vez condicionam os valores dos *ratings* comerciais.

I. Caracterização da Evolução da Variável *TVR_Prog%*

Verificámos em análises posteriores uma considerável correlação desta variável face aos valores de *ratings*, apresentando os modelos de Regressão Linear simples e os modelos de Regressão Linear multivariados simbólicos valores consideráveis dos coeficientes de determinação. Consequentemente, a previsão deste indicador apresenta uma grande relevância.

Com base nas análises gráficas das ST verificamos que esta variável apresenta uma evolução que difere significativamente entre os dias laborais e o fim-de-semana, existindo alguma semelhança no decorrer de cada um destes períodos. Uma vez que os seus valores dizem respeito a valores programáticos existe desfasamento entre os instantes temporais a que se referem as observações de cada série temporal, logo não é viável proceder ao seu

clustering sem previamente alinhar as séries. Após a análise de *TVR_Período%* verificar-se-á se é necessário proceder às referidas análises.

Em dias laborais existe um primeiro pico de exposição ao *media*, que corresponde ao espaço informativo da tarde e a partir do qual se verifica uma acentuada redução nestes índices até ao programa com início por volta das 19H. Um segundo pico de audiências, muito superior ao da hora de almoço pode ser associado ao período que decorre desde o espaço informativo da noite até ao programa com início no período entre as 22H e as 23H e a partir do qual se verifica uma vez mais uma acentuada redução nesta variável (*Gráfico F.1.1.5.I.1*).

Ao fim-de-semana observa-se a mesma tendência, mas em geral com uma redução mais tardia dos valores de *TVR_Prog%* no período da noite, mantendo-se os valores mais elevados até ao programa com início às 23H e que se poderá prolongar até às 24H. No decorrer da tarde os seus valores também são mais elevados relativamente a períodos equivalentes dos dias laborais.

Existem algumas evoluções diárias que visivelmente se afastam da média e que correspondem a transmissões de programas com uma maior popularidade, como, por exemplo, determinados jogos de futebol ou a transmissões associadas a processos judiciais (como, por exemplo, Casa Pia), entre outros eventos e notícia de “última hora”. Para os vários anos considerados a repercussão dos eventos desportivos apresenta uma grande variabilidade.

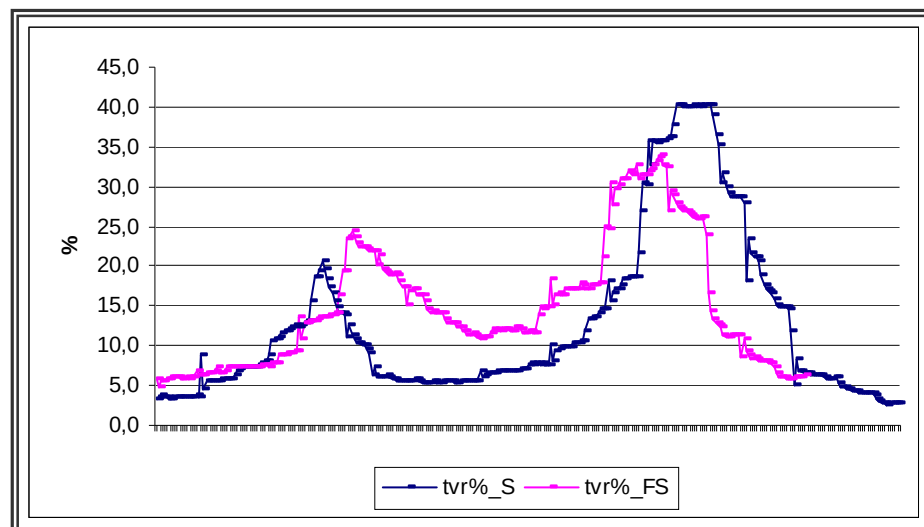


Gráfico F.1.1.5.I.1 – Evolução da variável *TVR_Prog%* associada ao período semanal e de fim-de-semana

Excluindo alguns programas que geram observações atípicas, aparentemente no decorrer dos cinco anos em que são superiores as alterações estruturais deste sector, isto é, até 2004, a evolução diária observada em *TVR_Prog%* apresenta alguma semelhança.

Desde que não se verifiquem significativas alterações na grelha de programas existe um padrão diário de evolução desta variável. Observamos alguma estabilidade de valores para períodos horários equivalentes. Para a avaliação e quantificação da variabilidade deste indicador recorre-se posteriormente aos valores associados a períodos horários.

Como foi referido anteriormente, com base em dados agregados de audiências não é viável poder garantir que esta aparente estabilidade de valores se deva a hábitos de consumo mais regulares para a maior parte do painel de audimetria. Podem existir flutuações individuais que se complementem e anulem possíveis oscilações existentes nos dados, existindo apenas alguma regularidade do ponto de vista agregado.

No entanto, provavelmente as selecções de alguns segmentos de telespectadores também podem estar condicionadas por parte de determinados factores estruturais, como, por exemplo, a disponibilidade, dependendo também o grau da sua influência das faixas horárias consideradas nas análises.

A evolução desta variável é independente do canal considerado, uma vez que corresponde ao global da exposição televisiva. No entanto, os canais podem-se diferenciar quanto ao seu contributo para a variabilidade observada neste indicador, em função do peso de cada segmento sociodemográfico no seu público-alvo e da estruturação das suas grelhas. A variabilidade deste indicador poderá estar relacionada com grupos específicos de telespectadores, nomeadamente os que apresentam menores “tempos vazios” e mais variáveis, e que consequentemente também se irá reflectir ao nível dos *ratings* e dos *shares* de audiências. Logo, a comparação da evolução deste indicador para os diversos segmentos também será posteriormente necessária.

Os dados para programas em concreto, abrangem o período horário associado à sua duração global ou parcial. Podemos calcular para cada um dos programas que apresenta vários episódios um intervalo de variação de *TVR_Prog%*, transformando os dados em variáveis simbólicas da tipologia intervalo. A sua amplitude reflecte, não só, a variação verificada no número de telespectadores disponíveis ao longo da semana por factores alheios à programação, como também será consequência das alterações dos períodos horários em que estes programas são transmitidos e das respectivas durações.

Verificamos que a variabilidade desta variável pode ser muito significativa, podendo alcançar os 24% se considerarmos os dados desde 2000 a 2008 e como horizonte temporal em cada ano os dois meses que decorrem desde 15 de Abril a 15 de Junho. A sua variabilidade será certamente superior à da variável *TVR_Período%* e em especial se considerarmos a durabilidade programática global face aos correspondentes valores para suportes temporais mais restritos. Consequentemente, os factores temporais e de disponibilidade do painel imperam, contribuindo em geral para a existência de uma reduzida lealdade programática e para com o meio televisivo e que se confirmará com base nas análises de duplicação das audiências.

A alteração dos horários de transmissão repercute-se de forma muito significativa neste indicador, logo as previsões têm que ser realizadas com base nos valores para faixas horárias. Quando as metodologias de previsão promovem a identificação de cenários que sejam mais semelhantes à observação em análise será necessário considerar apenas as

observações associadas a idênticas faixas horárias ou segmentos horários que resultem das análises de segmentação. Em geral a curto prazo as previsões devem ser realizadas para intervalos de tempo de 15 ou 30 minutos. Contudo, quando se pretende prever a médio e longo prazo deve-se considerar uma maior duração dos suportes horários e que devem estar associados ao modo de vida dos telespectadores no decorrer do dia e que podem ser identificados com base em análises de segmentação das séries temporais.

Recorrendo a algoritmos de *clustering* simbólicos verificamos que os programas com maiores intervalos de variabilidade correspondem em maior grau a determinadas faixas horárias, entre os quais se destacam, por exemplo, os espaços informativos e os programas transmitidos no decorrer da noite. Logo, nestes contextos a previsão com base em dados simbólicos que codifiquem os dados difusos são especialmente relevantes.

Em princípio não será o mesmo prever as audiências de programas transmitidos no decorrer do horário nobre ou no decorrer da noite, uma vez que no horário nobre coincidem os horários de início dos programas e no decorrer da noite, em geral o planeamento da grelha é realizado de modo a não fazer coincidir o início dos programas, evitando a “fuga” dos telespectadores para os canais concorrentes. Contudo, em determinados anos, como, por exemplo, em 2012 as grelhas da noite frequentemente estão estruturadas com uma sequência de horários de início dos programas muito semelhante entre os diversos canais generalistas, em especial se considerarmos os canais privados.

Verificamos que existe diferenciação no padrão de variabilidade desta variável em função dos programas e tipologias programáticas transmitidas. Numa tentativa de promover a identificação dos factores responsáveis pela maior variabilidade de valores foi possível retirar algumas conclusões. Privilegiaram-se os programas mais relevantes para as agências de planificação de meios do ponto de vista do seu valor comercial.

□ Os Espaços Informativos

Os espaços informativos da tarde apresentam menor variabilidade de valores face aos programas equivalentes da noite de Segunda a Sexta, uma vez que existe não só uma maior estabilidade nos seus horários de transmissão, como também, ao nível da estruturação da grelha em contra-programação e em princípio quanto aos factores de disponibilidade do painel.

Verificamos também que o indicador associado aos espaços informativos da tarde em dias laborais é menos variável face ao correspondente programa de fim-de-semana. Isto poderá ser atribuído à existência de um público de maior dimensão ao fim-de-semana e onde, conseqüentemente, mais facilmente ocorrem alterações nos seus padrões individuais de disponibilidade ou à influência de determinados programas. Ao fim-de-semana mais facilmente os tempos vazios são substituídos por actividades de lazer.

A alteração de um programa para diferentes períodos horários repercute-se significativamente no valor desta variável, especialmente se anteciparmos o seu início e

em especial para o canal estatal RTP1 onde aparentemente existe uma selecção programática mais selectiva. O público dos canais RTP aparentemente é o que influencia em maior grau a variabilidade observada neste indicador, por existir uma menor fidelização do seu público para com o meio televisivo.

Por exemplo, atendendo aos espaços informativos da tarde, os programas que apresentam menores variações de *TVR_Prog%* são o programa semanal da RTP1 com 2,3% de variabilidade de valores (10% ao fim-de-semana), o programa semanal da TVI com 2,7% (6,6% ao fim-de-semana) e o programa semanal da SIC com 2% (5,7% ao fim-de-semana). Os valores de variabilidade atendendo ao que é considerado como aceitável no meio televisivo não são muito significativos no decorrer dos dias laborais. No mercado Norte Americano esta margem é de aproximadamente 10%. Analisando os espaços informativos da tarde ao fim-de-semana verificamos que a variabilidade de valores de *TVR_Prog%* pode ser essencialmente atribuída à influência de determinados programas, por exemplo, a Casamentos Reais e às transmissões de Fórmula 1 por parte da RTP1. Se ignorarmos estes programas reduzimos a variabilidade desta variável para 2,3% para o canal SIC (5,7% era o valor original), para a TVI ficamos com apenas 2,3% de variabilidade (6,3% era o valor original) e para a RTP1 ficamos com 2,7% face aos 10% iniciais. Uma vez mais inclusive em períodos de lazer a variabilidade poderá ser aceitável se considerarmos uma grelha de programação já estabelecida e um horizonte temporal não influenciado pela sazonalidade do consumo televisivo.

As análises de previsão devem ser realizadas por horizontes temporais que apresentem uma maior uniformidade relativamente à sazonalidade presente no consumo televisivo.

Consequentemente, a repercussão de determinados programas pode ser muito significativa e será função de factores dificilmente quantificáveis, como, por exemplo, as preferências individuais e do agregado familiar, a qualidade e notoriedade da programação, o conhecimento da grelha, factores de disponibilidade, entre outras variáveis estruturais e individuais. Nestes contextos em geral recorre-se às sondagens de intenção de ver, apesar das suas fortes limitações. No *capítulo V* apresenta-se um modelo estrutural de exposição ao *media* e no *capítulo VI* sugere-se uma abordagem de construção de planos de *media* que se adequa a estes contextos.

Ao fim-de-semana mais facilmente poderá existir um acréscimo de valores de *TVR_Prog%*, uma vez que a maior parte do painel poderá ter uma maior disponibilidade. No entanto, este não será o único factor a influenciar os índices de audiências e a sua repercussão não será obrigatoriamente positiva. Por, exemplo, a transmissão dos programas de Fórmula 1 em geral antecipa o horário do espaço informativo da tarde, o que provoca uma diminuição nos valores de *TVR_Prog%* associados a este programa e a um acréscimo na sua variabilidade. Nestas situações os canais privados saem beneficiados por transmitirem os seus programas simultaneamente com as transmissões de Fórmula 1, incrementando os valores de *TVR_Prog%* dos seus programas.

Quanto aos espaços informativos da noite, o espaço informativo da noite semanal da RTP1 atinge valores de 25,3% de variabilidade de *TVR_Prog%* (18,4% ao FS), o programa equivalente da TVI de 15% (9,8% ao FS) e o programa da SIC de 14,2% (9,7% ao FS). Estes valores resultam de considerar a totalidade da amostra, não se excluindo programas mais pontuais que possam favorecer alterações na evolução diária desta variável. Em relação ao valor associado ao espaço informativo da noite da RTP1 a sua grande variabilidade resulta da transmissão de eventos desportivos antes da sua emissão e, consequentemente na alteração do seu horário para mais tarde, o que impulsiona de forma muito significativa as taxas de audiência e que também se poderá repercutir nos índices associados aos canais concorrentes. Nas previsões será necessário ter em atenção as faixas horárias de transmissão e as tipologias concorrentes, mas também as adjacentes.

Apenas dois jogos de futebol envolvendo o FC Porto provocam um acréscimo mais significativo na variável *TVR_Prog%*. Neste contexto o espaço informativo da noite da RTP1 também beneficia dos efeitos de herança existentes e por não existir em contra-programação a transmissão de espaços informativos. Alguns eventos captam um maior número de telespectadores no seu decorrer, mas são incapazes por efeito de inércia de os transportar para os programas adjacentes, levando o seu público a ligar os canais concorrentes ou a desligar o televisor. A previsão da repercussão da transmissão de eventos desportivos ou de outros programas de maior notoriedade apresenta uma grande complexidade.

Para o canal TVI se excluirmos o espaço informativo da noite que apresenta como contra-programação os jogos de futebol, a variabilidade fica reduzida a 7,7% face aos 15% iniciais e para a SIC reduz-se para 7,4% face aos 14,2% iniciais. Existe uma redução para metade da variabilidade observada nestes indicadores.

Considerando a totalidade dos programas, verificamos que ao fim-de-semana é menor a variabilidade dos valores de *TVR_Prog%* para os espaços informativos da noite e excluindo programas que geram observações atípicas mantém-se este resultado.

Verificamos que os maiores valores de variabilidade de *TVR_Prog%* resultam essencialmente das contínuas alterações que se verificam nos horários de transmissão, e por vezes na contra-programação dos programas. Portanto, pontualmente a variabilidade pode ser muito considerável, especialmente para programas com horários de início mais variáveis do período da noite, horário nobre e pré-prime-time e que apresentem em contra-programação determinados jogos de futebol.

Os valores médios desta variável seguem em alguns períodos uma evolução exponencial ou polinomial, incrementando ou decrescendo fortemente os seus valores em curtos espaços de tempo para a maior parte da emissão diária. Consequentemente, pequenas oscilações nos horários de transmissão dos programas podem afectar significativamente estes valores e provavelmente também os valores dos *ratings*.

As oscilações nos valores dos *ratings* serão tanto menores quanto mais predominarem os comportamentos de lealdade tipológica, programática e ao próprio canal. Os espaços informativos funcionam como pilares da grelha de programas, sendo os programas directamente adjacentes geralmente os mais favorecidos e os programas mais afastados são os mais penalizados.

A transmissão de programas pontualmente mais “atractivos” para a maior parte do público, como, por exemplo, os jogos de futebol será outro factor a considerar. No entanto, os eventos desportivos não se podem equiparar quanto à sua repercussão nesta variável e os restantes eventos, como, por exemplo, o Festival Eurovisão, galas, entre outros, ficam bastante aquém na sua influência para o período em análise.

Não podemos ignorar totalmente o conteúdo programático e conduzir as análises apenas com base em tipologias programáticas que não deixam de ser categorizações impostas pela indústria televisiva e que nem sempre correspondem à segmentação efectuada implicitamente pelos telespectadores.

A utilização de análises MDS e de Correspondência viabiliza a associação de programas aos telespectadores. Analisando os valores históricos associados a diferentes cenários de contra-programação, envolvendo esta tipologia de programas talvez se possa calcular um coeficiente que quantifique a sua influência nos valores médios observados e que são função das faixas horárias consideradas.

Contudo, apenas podemos considerar o passado recente e dificilmente se vão encontrar um significativo número de observações, o que condicionará a precisão das estimativas realizadas. Se considerarmos um período mais alargado verificamos que a variabilidade de *TVR%* sobe consideravelmente. A previsão de índices de audiências programáticos está consequentemente fortemente dependente do período horário em que um programa é transmitido, e por vezes da repercussão da contra-programação e da programação adjacente. A tipologia programática é insuficiente para prever possíveis acréscimos e decréscimos neste indicador decorrente da introdução de programas mais populares, por exemplo, dos eventos desportivos que na realidade não são equiparáveis na sua repercussão. Variáveis mais discriminatórias dos conteúdos podem ter que ser consideradas.

Contudo, no *Capítulo V* demonstra-se que algumas variáveis que não são previsíveis influenciam a exposição, e que por vezes são muito contraditórias as avaliações a respeito da sua influência. Os períodos horários influenciam o grau de variabilidade que se poderá verificar nos valores de *TVR%* e a composição sociodemográfica do painel e a estruturação da grelha de programação para esses períodos influencia os valores dos *shares* de audiências e de *ratings* em função das tipologias programáticas e dos programas transmitidos.

□ Os Espaços Destinados aos Filmes

Considerando os espaços destinados à transmissão de filmes, por exemplo, da SIC, os maiores valores de variabilidade também correspondem aos filmes que tiveram como

contra-programação, os eventos desportivos. Excluindo-os da amostra a variabilidade apresenta o valor máximo de 2,5%, podendo ascender a 9,4% caso contrário. Logo a variabilidade é reduzida, apesar desta tipologia de programas ser considerada como uma das que coloca maiores dificuldades, uma vez que em geral também são transmitidos filmes em contra-programação e os canais não divulgam com grande antecedência os conteúdos que vão ser transmitidos, numa tentativa de, assim, maximizar os seus *shares* de audiências.

Os filmes que apresentam uma menor variabilidade de valores correspondem, quer a programas de fim-de-semana, quer dos dias laborais do final da noite.

Os filmes que apresentam maiores valores são os transmitidos aos Domingos com início entre as 17H-18.15H (14,9%), transmitidos ao Sábado, por volta das 23H (9,4%) e transmitidos ao Sábado por volta das 16.30H-18.30H (7,7%). Neste contexto correspondem essencialmente a programas de fim-de-semana, mas não só e do decorrer da tarde e da noite. O maior valor corresponde ao espaço que teve como contra-programação a transmissão de eventos desportivos por parte do canal estatal RTP1 em determinados dias. Logo, os programas desportivos são os responsáveis pela variabilidade observada em detrimento da eventual irregularidade nos padrões de exposição televisiva.

Em relação aos filmes da TVI também apresentam maiores valores de *TVR_Prog%* e uma maior variabilidade resultante das oscilações horárias e da introdução de programas de maior popularidade. Os filmes da noite para este canal apresentam uma maior estabilidade (<3,2%) em relação aos filmes da tarde. Para a RTP1 os menores intervalos de *TVR_Prog%* estão associados aos filmes de Sábado à tarde (0,3%) e a filmes transmitidos em dias laborais à noite (0,1%) e os maiores valores aos filmes de Domingo transmitidos ao final da tarde (21,8%). Os maiores valores de variabilidade resultam da transmissão de programas de maior popularidade e das oscilações horárias dos programas.

□ Os Programas do Horário Nobre

Em relação aos restantes programas, os maiores valores de variabilidade para o canal SIC e TVI correspondem às telenovelas do horário nobre ou do início da noite. Para o canal estatal RTP1 os maiores valores correspondem aos concursos de pré-prime-time e da noite. As faixas horárias a partir das 19H são sem dúvida as que apresentam uma maior variabilidade de valores. Para o canal SIC a maior variabilidade de valores, com um valor aproximado de 10%, corresponde às duas principais telenovelas da noite da semana e à telenovela de pré-prime-time (13,6%), seguindo-se a última telenovela da noite. Para o canal TVI os maiores valores correspondem à principal telenovela da noite no decorrer de toda a semana, atingindo para o período em análise o valor de 17,5% (S) e à telenovela Morangos com Açúcar (FS) com um valor muito elevado para o período em análise (31%).

Por último, para o canal estatal os valores maximizam-se para o concurso “o Preço Certo” e para o programa de humor de fim-de-semana da noite, seguindo-se os programas equivalentes em dias laborais.

Portanto, pontualmente a variabilidade pode ser muito considerável, especialmente para programas com horários de início mais variáveis do período da noite, horário nobre e pré-prime-time e que apresentem como contra-programação determinados jogos de futebol.

Os valores de *TVR_Prog%* em função das durações programáticas parciais podem ser menos variáveis, uma vez que correspondem a intervalos de tempo mais restritos. Do ponto de vista comercial é mais relevante a análise das audiências programáticas parciais, uma vez que os espaços comerciais estão associados a partes de programas e são tarifadas de forma diferenciada, ainda que sejam transmitidos no decorrer do mesmo programa.

Com base na análise dos valores parciais de *TVR_Prog%* verificamos que alguns programas apresentam valores de variabilidade bastante reduzidos, e por vezes a variabilidade global pode ser atribuída apenas a determinadas partes programáticas. Atendendo aos filmes do período da noite é notório o decréscimo verificado no seu decorrer quanto à variabilidade e dimensão de *TVR_Prog%* (P).

Os períodos com maiores níveis de audiências também são os mais variáveis, o que se poderá atribuir em maior grau aos segmentos que apresentam uma menor regularidade na sua exposição televisiva. Quanto aos filmes da tarde após a conclusão do espaço informativo a variabilidade vai decrescendo gradualmente

invertendo-se esta tendência com uma maior proximidade face ao espaço informativo da noite. Os filmes da RTP1 e em especial os que são transmitidos no período da noite podem apresentar uma grande amplitude de valores.

Nas análises de previsão será necessário ter em atenção que as audiências de semana para semana variam de forma diferenciada em função dos programas e da sua localização na grelha. Existem programas que em determinadas faixas horárias e canais apresentam e incutem uma maior variância aos dados. Outros pelo contrário apresentam uma variabilidade muito menor, o que condiciona os instrumentos de previsão mais adequados.

II. Caracterização da evolução da variável TVR_Período%

Com base nas análises realizadas identificámos quais são os factores com uma maior repercussão na variabilidade dos dados e a influência decorrente das alterações que se verificam nos horários de transmissão dos programas. Os valores indicam a existência de um reduzido envolvimento com o meio televisivo e com os conteúdos programáticos normalmente transmitidos, para além de uma elevada capacidade do *media* para atrair um significativo número de telespectadores quando os conteúdos transmitidos o justificam. Consequentemente, as motivações dos telespectadores apenas pontualmente aparentam ser instrumentais (Cooper & Tang, 2009).

Neste apêndice pretende-se a análise das taxas médias de audiências para o total da TV em função das faixas horárias, considerando intervalos de 30 ou 15 minutos.

Estes períodos correspondem ao horizonte temporal que é utilizado na prática quando se pretende realizar previsões a curto prazo.

Atendendo às semelhanças existentes na evolução desta variável para os vários anos considerados, interessa avaliar até que ponto é possível identificar períodos horários com uma evolução mais comparável. Nos períodos que apresentem uma menor variabilidade da variável em estudo, mais facilmente se poderá prever o seu valor, provavelmente recorrendo a metodologias univariadas de análise de ST. Por outro lado, as faixas horárias onde se verifique uma evolução mais variável devem ser objecto de uma análise posterior mais exhaustiva.

As estratégias de programação e de contra-programação, assim como a composição sociodemográfica da audiência é variável no decorrer do dia. Consequentemente uma vez que está principalmente em causa o desenvolvimento de modelos explicativos de *ratings*, e não univariados, é necessário segmentar a emissão diária em períodos que apresentem tendências similares e que posteriormente devem ser alvo de análises diferenciadas. Devemos verificar se a classificação adoptada tradicionalmente pelo sector quanto à definição das faixas horárias corresponde à que resulta da segmentação dos dados, atendendo à sua variabilidade.

A segmentação dos dados também permite uma representação das séries originais que viabiliza se necessário a posterior aplicação de metodologias do âmbito do “Rule Discovery”, “Novelty Detection” ou do “Motif Discovery” que podem eventualmente contribuir para a previsão deste indicador.

Quanto menor for a variabilidade inter-anual para idênticos períodos horários, mais razoável será podermos concluir que estamos perante um painel que apresenta um consumo essencialmente influenciado por condicionantes temporais e estruturais. Contudo, o comportamento agregado, nem sempre apresenta uma correspondência directa com o que se verifica a nível individual. Se verificarmos a existência de uma considerável estabilidade destes valores, denotando uma eventual maior passividade da audiência, a abordagem univariada para a sua previsão poderá ser equacionada, uma vez que os seus valores serão essencialmente função de factores temporais e em geral esta corresponde à abordagem mais utilizada na prática.

A partir dos dados de *TVR_Prog%* associados a programas procedeu-se ao cálculo de uma estimativa destes valores para períodos horários. Obtemos assim uma série de valores discretos, equitativamente espaçados no tempo. Para a construção das bases de dados considerou-se intervalos de meia hora utilizando valores globais e parciais de *TVR_Prog%*. Consideram-se para além das séries temporais de valores de *TVR%* associados a períodos horários, os intervalos de variabilidade desta variável para cada período e que se traduz em uma base de dados simbólica. As datas em análise coincidem para os vários anos considerados evitando-se assim diferenças que possam ser atribuídas à sazonalidade do consumo. Apesar de se tentar minimizar as desigualdades existentes nas condições subjacentes que possam influenciar o consumo televisivo, estas não podem ser totalmente excluídas.

Meras alterações climáticas poderá ser suficiente para se alterarem os padrões de consumo televisivo entre os vários anos considerados. As alterações que se possam verificar na grelha, ao longo dos anos, também podem influenciar as diferenças identificadas, sobretudo as que decorrem da transmissão de eventos desportivos. Obviamente que dificilmente se mantêm os factores pessoais e individuais que condicionam os padrões de disponibilidade e de predisposição para ver TV.

IIa. A evolução de *TVR_Periodo%*

Observamos um padrão diário de evolução deste indicador, o qual se diferencia entre o fim-de-semana e os dias laborais. As grandes semelhanças existentes na evolução diária desta variável mantêm-se inclusive para os anos mais próximos. Até certo ponto existe alguma estabilidade nos seus valores para idênticos períodos, tornando este indicador independente da programação e a sua variabilidade pode ser atribuída em maior grau a factores de disponibilidade/ vontade do painel. Programas de grande popularidade podem alterar significativamente a evolução desta variável. Nos gráficos que ilustram a evolução média desta variável observamos dois picos de audiências que correspondem aos espaços informativos e ao horizonte temporal denominado de horário nobre. A restante emissão diária subdivide-se em diferentes períodos com tendências ascendentes ou descendentes do tipo polinomial e linear, estando os valores em determinado instante obviamente correlacionados com os adjacentes, em parte também pelos efeitos de herança existentes entre os programas adjacentes

A análise dos índices de duplicação consta dos capítulos subsequentes, no entanto segundo a bibliografia internacional estes valores não costumam ser muito elevados. Este factor poderá ser uma das causas para a inexistência de um significativo número de períodos onde se possa observar uma maior estabilidade de valores de *ratings*. A associação entre a evolução dos *ratings* e a duplicação primária das audiências consta do *Apêndice F.4*.

Verificamos que atendendo à evolução média de *TVR_Período%* para os dias laborais, assistimos no decorrer da manhã até ao espaço informativo da tarde e a partir das 16.30H a uma tendência ascendente polinomial e após os dois espaços informativos um decréscimo polinomial no início da tarde e linear no decorrer da noite. Apenas entre as 13H e as 14H e entre as 20.30H e as 22.30H existe uma maior estabilidade na dimensão desta variável, correspondendo estes períodos ao espaço informativo da tarde e à segunda parte do espaço informativo da noite e seus programas sucessores. Correspondem às faixas horárias onde após um dia de trabalho as pessoas recorrem à TV por esta ser um veículo de informação, e posteriormente assistem a um reduzido número de programas de ficção. O abandono no decorrer da noite a nível médio processa-se linearmente. A emissão diária a nível médio poderia ser subdividida nos seguintes segmentos horários: desde o início da emissão até 13H, 13H-14H, 14H-16.30H, 16.30H-20H, 20H-22H e 22H até final da emissão. Não corresponde à divisão que *Fernandes (2001)* apresenta como sendo adoptada na indústria televisiva. Ao fim-de-semana e a nível médio a emissão diária também pode ser subdividida em quatro períodos de transição das audiências, com

tendências polinomiais (7.30H-13H; 13.30H-16.30H; 16.30H-20H) e linear (22.30H-FIM). A evolução de *TVR%* em determinados períodos é menos acentuada e o patamar de audiências médio é meia hora mais longo no período da noite (em dimensão ligeiramente inferior aos dos dias laborais, 2-3%), mas idêntico ao semanal à hora do almoço (em dimensão superior ao semanal em 5%). Uma vez mais por volta das 16.30H verifica-se um ponto de inflexão na evolução de *TVR_Período%*. Portanto, sem dúvida que os períodos horários se podem diferenciar quanto às taxas de crescimento e decréscimo, estando os valores desta variável em determinado instante correlacionados com os adjacentes. A proximidade de valores nas tendências médias não é, no entanto, condição suficiente para a variabilidade de valores, ano após ano, ser reduzida para idênticos períodos horários.

IIb. Variabilidade da variável *TVR_Período%*

A maior disponibilidade do público ao fim-de-semana gera uma maior variabilidade na evolução desta variável ao longo de todo o dia até ao período que antecede o espaço informativo da noite, sensivelmente as 19.30H, após o qual existe uma audiência total mais estável. Ao fim-de-semana a existência de uma audiência sem condicionantes laborais reduz a variabilidade observada ao nível dos períodos de maiores audiências no período da noite, mas a partir das 22.30H os valores tendem a equiparar-se, ainda que o período semanal apresente maiores valores médios a partir das 22.30H.

Uma vez que ao longo da semana as pessoas que não exercem uma actividade profissional são as que podem assistir maioritariamente à televisão até às 19H os padrões de consumo semanais apresentam uma maior estabilidade, ou seja, existe uma audiência mais regularmente disponível.

Os valores mínimos semanais de *TVR%_Período* são inferiores aos de fim-de-semana até às 20H, e posteriormente equivalentes ou ligeiramente superiores. Os valores máximos semanais de *TVR_Período%* são inferiores até às 19H e posteriormente inverte-se esta tendência.

Verificamos que durante os dias laborais a variabilidade dos valores é máxima entre as 19H e as 21.30H (13% a 20%) considerando a totalidade da grelha, mas para as restantes faixas horárias, os valores são inferiores aos 6%. Ao fim-de-semana, excluindo o período que decorre entre as 17H e as 20H, os valores de variabilidade ficam aquém dos 10%. Em geral a variabilidade de valores é superior ao fim-de-semana.

Ao fim-de-semana os valores médios, são superiores à hora do almoço, mas inferiores no período de maiores audiências da noite, relativamente aos valores dos dias laborais. Por exemplo, para os anos de maior variabilidade desta variável os valores são no máximo de 15% para os dias laborais e de 10% ao fim-de-semana.

IIC. Evolução Inter-Anual de *TVR_Período%*

Comparando a evolução inter-anual verificamos que o diferencial dos valores de *TVR_Período%* varia em função da distância inter-anual, faixas horárias consideradas e consoante nos referimos a dias laborais ou ao fim-de-semana. As faixas horárias não são equiparáveis quanto à sua propensão para se verificarem alterações nos padrões de disponibilidade do painel.

A hora do almoço apresenta valores de variabilidade quase duplos aos da noite para os dias laborais. O fim-de-semana e a hora de jantar dos dias laborais apresentam uma maior estabilidade de valores entre os vários anos considerados, sendo estes também os horizontes temporais mais relevantes do ponto de vista comercial.

Portanto, verificamos que a nível global existe no decorrer dos anos uma maior estabilidade nos padrões agregados da exposição em períodos de lazer do que em dias laborais. Os períodos que correspondem a uma actividade laboral apresentam um comportamento inter-anual mais diferenciado, apesar da exposição televisiva poder estar associada em maior grau a quem não exerce uma actividade profissional, ou pelo menos por conta de terceiros.

As manhãs diferenciam-se por apresentarem uma variabilidade superior ao fim-de-semana, mas todas as faixas horárias apresentam uma menor variabilidade inter-anual para os períodos de lazer (*Tabelas F.1.1.5.II.1 e F.1.1.5.II.2*).

Os dados para programas em concreto, abrangem o período horário associado à sua duração global ou parcial. Podemos calcular para cada um dos programas que apresenta vários episódios um intervalo de variação de *TVR_Prog%*, transformando os dados em variáveis simbólicas da tipologia intervalo. A sua amplitude reflecte, não só, a variação verificada no número de telespectadores disponíveis ao longo da semana por factores alheios à programação, como também será consequência das alterações dos períodos horários em que estes programas são transmitidos e das respectivas durações.

Verificamos que a variabilidade desta variável pode ser muito significativa, podendo alcançar os 24% se considerarmos os dados desde 2000 a 2008 e como horizonte temporal em cada ano os dois meses que decorrem desde 15 de Abril a 15 de Junho. A sua variabilidade será certamente superior à da variável *TVR_Período%* e em especial se considerarmos a durabilidade programática global face aos correspondentes valores para suportes temporais mais restritos. Consequentemente, os factores temporais e de disponibilidade do painel imperam, contribuindo em geral para a existência de uma reduzida lealdade programática e para com o meio televisivo e que se confirmará com base nas análises de duplicação das audiências.

A alteração dos horários de transmissão repercute-se de forma muito significativa neste indicador, logo as previsões têm que ser realizadas com base nos valores para faixas horárias. Quando as metodologias de previsão promovem a identificação de cenários que sejam mais semelhantes à observação em análise será necessário considerar apenas as

observações associadas a idênticas faixas horárias ou segmentos horários que resultem das análises de segmentação. Em geral a curto prazo as previsões devem ser realizadas para intervalos de tempo de 15 ou 30 minutos. Contudo, quando se pretende prever a médio e longo prazo deve-se considerar uma maior duração dos suportes horários e que devem estar associados ao modo de vida dos telespectadores no decorrer do dia e que podem ser identificados com base em análises de segmentação das séries temporais.

Recorrendo a algoritmos de *clustering* simbólicos verificamos que os programas com maiores intervalos de variabilidade correspondem em maior grau a determinadas faixas horárias, entre os quais se destacam, por exemplo, os espaços informativos e os programas transmitidos no decorrer da noite. Logo, nestes contextos a previsão com base em dados simbólicos que codifiquem os dados difusos são especialmente relevantes.

Em princípio não será o mesmo prever as audiências de programas transmitidos no decorrer do horário nobre ou no decorrer da noite, uma vez que no horário nobre coincidem os horários de início dos programas e no decorrer da noite, em geral o planeamento da grelha é realizado de modo a não fazer coincidir o início dos programas, evitando a “fuga” dos telespectadores para os canais concorrentes. Contudo, em determinados anos, como, por exemplo, em 2012 as grelhas da noite frequentemente estão estruturadas com uma sequência de horários de início dos programas muito semelhante entre os diversos canais generalistas, em especial se considerarmos os canais privados.

Verificamos que existe diferenciação no padrão de variabilidade desta variável em função dos programas e tipologias programáticas transmitidas. Numa tentativa de promover a identificação dos factores responsáveis pela maior variabilidade de valores foi possível retirar algumas conclusões. Privilegiaram-se os programas mais relevantes para as agências de planificação de meios do ponto de vista do seu valor comercial.

□ Os Espaços Informativos

Os espaços informativos da tarde apresentam menor variabilidade de valores face aos programas equivalentes da noite de Segunda a Sexta, uma vez que existe não só uma maior estabilidade nos seus horários de transmissão, como também, ao nível da estruturação da grelha em contra-programação e em princípio quanto aos factores de disponibilidade do painel.

Verificamos também que o indicador associado aos espaços informativos da tarde em dias laborais é menos variável face ao correspondente programa de fim-de-semana. Isto poderá ser atribuído à existência de um público de maior dimensão ao fim-de-semana e onde, conseqüentemente, mais facilmente ocorrem alterações nos seus padrões individuais de disponibilidade ou à influência de determinados programas. Ao fim-de-semana mais facilmente os tempos vazios são substituídos por actividades de lazer.

A alteração de um programa para diferentes períodos horários repercute-se significativamente no valor desta variável, especialmente se anteciparmos o seu início e

em especial para o canal estatal RTP1 onde aparentemente existe uma selecção programática mais selectiva. O público dos canais RTP aparentemente é o que influencia em maior grau a variabilidade observada neste indicador, por existir uma menor fidelização do seu público para com o meio televisivo.

Por exemplo, atendendo aos espaços informativos da tarde, os programas que apresentam menores variações de *TVR_Prog%* são o programa semanal da RTP1 com 2,3% de variabilidade de valores (10% ao fim-de-semana), o programa semanal da TVI com 2,7% (6,6% ao fim-de-semana) e o programa semanal da SIC com 2% (5,7% ao fim-de-semana). Os valores de variabilidade atendendo ao que é considerado como aceitável no meio televisivo não são muito significativos no decorrer dos dias laborais. No mercado Norte Americano esta margem é de aproximadamente 10%. Analisando os espaços informativos da tarde ao fim-de-semana verificamos que a variabilidade de valores de *TVR_Prog%* pode ser essencialmente atribuída à influência de determinados programas, por exemplo, a Casamentos Reais e às transmissões de Fórmula 1 por parte da RTP1. Se ignorarmos estes programas reduzimos a variabilidade desta variável para 2,3% para o canal SIC (5,7% era o valor original), para a TVI ficamos com apenas 2,3% de variabilidade (6,3% era o valor original) e para a RTP1 ficamos com 2,7% face aos 10% iniciais. Uma vez mais inclusive em períodos de lazer a variabilidade poderá ser aceitável se considerarmos uma grelha de programação já estabelecida e um horizonte temporal não influenciado pela sazonalidade do consumo televisivo.

As análises de previsão devem ser realizadas por horizontes temporais que apresentem uma maior uniformidade relativamente à sazonalidade presente no consumo televisivo.

Consequentemente, a repercussão de determinados programas pode ser muito significativa e será função de factores dificilmente quantificáveis, como, por exemplo, as preferências individuais e do agregado familiar, a qualidade e notoriedade da programação, o conhecimento da grelha, factores de disponibilidade, entre outras variáveis estruturais e individuais. Nestes contextos em geral recorre-se às sondagens de intenção de ver, apesar das suas fortes limitações. No *capítulo V* apresenta-se um modelo estrutural de exposição ao *media* e no *capítulo VI* sugere-se uma abordagem de construção de planos de *media* que se adequa a estes contextos.

Ao fim-de-semana mais facilmente poderá existir um acréscimo de valores de *TVR_Prog%*, uma vez que a maior parte do painel poderá ter uma maior disponibilidade. No entanto, este não será o único factor a influenciar os índices de audiências e a sua repercussão não será obrigatoriamente positiva. Por, exemplo, a transmissão dos programas de Fórmula 1 em geral antecipa o horário do espaço informativo da tarde, o que provoca uma diminuição nos valores de *TVR_Prog%* associados a este programa e a um acréscimo na sua variabilidade. Nestas situações os canais privados saem beneficiados por transmitirem os seus programas simultaneamente com as transmissões de Fórmula 1, incrementando os valores de *TVR_Prog%* dos seus programas.

Quanto aos espaços informativos da noite, o espaço informativo da noite semanal da RTP1 atinge valores de 25,3% de variabilidade de *TVR_Prog%* (18,4% ao FS), o programa equivalente da TVI de 15% (9,8% ao FS) e o programa da SIC de 14,2% (9,7% ao FS). Estes valores resultam de considerar a totalidade da amostra, não se excluindo programas mais pontuais que possam favorecer alterações na evolução diária desta variável. Em relação ao valor associado ao espaço informativo da noite da RTP1 a sua grande variabilidade resulta da transmissão de eventos desportivos antes da sua emissão e, consequentemente na alteração do seu horário para mais tarde, o que impulsiona de forma muito significativa as taxas de audiência e que também se poderá repercutir nos índices associados aos canais concorrentes. Nas previsões será necessário ter em atenção as faixas horárias de transmissão e as tipologias concorrentes, mas também as adjacentes.

Apenas dois jogos de futebol envolvendo o FC Porto provocam um acréscimo mais significativo na variável *TVR_Prog%*. Neste contexto o espaço informativo da noite da RTP1 também beneficia dos efeitos de herança existentes e por não existir em contra-programação a transmissão de espaços informativos. Alguns eventos captam um maior número de telespectadores no seu decorrer, mas são incapazes por efeito de inércia de os transportar para os programas adjacentes, levando o seu público a ligar os canais concorrentes ou a desligar o televisor. A previsão da repercussão da transmissão de eventos desportivos ou de outros programas de maior notoriedade apresenta uma grande complexidade.

Para o canal TVI se excluirmos o espaço informativo da noite que apresenta como contra-programação os jogos de futebol, a variabilidade fica reduzida a 7,7% face aos 15% iniciais e para a SIC reduz-se para 7,4% face aos 14,2% iniciais. Existe uma redução para metade da variabilidade observada nestes indicadores.

Considerando a totalidade dos programas, verificamos que ao fim-de-semana é menor a variabilidade dos valores de *TVR_Prog%* para os espaços informativos da noite e excluindo programas que geram observações atípicas mantém-se este resultado.

Verificamos que os maiores valores de variabilidade de *TVR_Prog%* resultam essencialmente das contínuas alterações que se verificam nos horários de transmissão, e por vezes na contra-programação dos programas. Portanto, pontualmente a variabilidade pode ser muito considerável, especialmente para programas com horários de início mais variáveis do período da noite, horário nobre e pré-prime-time e que apresentem em contra-programação determinados jogos de futebol.

Os valores médios desta variável seguem em alguns períodos uma evolução exponencial ou polinomial, incrementando ou decrescendo fortemente os seus valores em curtos espaços de tempo para a maior parte da emissão diária. Consequentemente, pequenas oscilações nos horários de transmissão dos programas podem afectar significativamente estes valores e provavelmente também os valores dos *ratings*.

As oscilações nos valores dos *ratings* serão tanto menores quanto mais predominarem os comportamentos de lealdade tipológica, programática e ao próprio canal. Os espaços informativos funcionam como pilares da grelha de programas, sendo os programas directamente adjacentes geralmente os mais favorecidos e os programas mais afastados são os mais penalizados.

A transmissão de programas pontualmente mais “atractivos” para a maior parte do público, como, por exemplo, os jogos de futebol será outro factor a considerar. No entanto, os eventos desportivos não se podem equiparar quanto à sua repercussão nesta variável e os restantes eventos, como, por exemplo, o Festival Eurovisão, galas, entre outros, ficam bastante aquém na sua influência para o período em análise.

Não podemos ignorar totalmente o conteúdo programático e conduzir as análises apenas com base em tipologias programáticas que não deixam de ser categorizações impostas pela indústria televisiva e que nem sempre correspondem à segmentação efectuada implicitamente pelos telespectadores.

A utilização de análises MDS e de Correspondência viabiliza a associação de programas aos telespectadores. Analisando os valores históricos associados a diferentes cenários de contra-programação, envolvendo esta tipologia de programas talvez se possa calcular um coeficiente que quantifique a sua influência nos valores médios observados e que são função das faixas horárias consideradas.

Contudo, apenas podemos considerar o passado recente e dificilmente se vão encontrar um significativo número de observações, o que condicionará a precisão das estimativas realizadas. Se considerarmos um período mais alargado verificamos que a variabilidade de *TVR%* sobe consideravelmente. A previsão de índices de audiências programáticos está consequentemente fortemente dependente do período horário em que um programa é transmitido, e por vezes da repercussão da contra-programação e da programação adjacente. A tipologia programática é insuficiente para prever possíveis acréscimos e decréscimos neste indicador decorrente da introdução de programas mais populares, por exemplo, dos eventos desportivos que na realidade não são equiparáveis na sua repercussão. Variáveis mais discriminatórias dos conteúdos podem ter que ser consideradas.

Contudo, no *Capítulo V* demonstra-se que algumas variáveis que não são previsíveis influenciam a exposição, e que por vezes são muito contraditórias as avaliações a respeito da sua influência. Os períodos horários influenciam o grau de variabilidade que se poderá verificar nos valores de *TVR%* e a composição sociodemográfica do painel e a estruturação da grelha de programação para esses períodos influencia os valores dos *shares* de audiências e de *ratings* em função das tipologias programáticas e dos programas transmitidos.

□ Os Espaços Destinados aos Filmes

Considerando os espaços destinados à transmissão de filmes, por exemplo, da SIC, os maiores valores de variabilidade também correspondem aos filmes que tiveram como

contra-programação, os eventos desportivos. Excluindo-os da amostra a variabilidade apresenta o valor máximo de 2,5%, podendo ascender a 9,4% caso contrário. Logo a variabilidade é reduzida, apesar desta tipologia de programas ser considerada como uma das que coloca maiores dificuldades, uma vez que em geral também são transmitidos filmes em contra-programação e os canais não divulgam com grande antecedência os conteúdos que vão ser transmitidos, numa tentativa de, assim, maximizar os seus *shares* de audiências.

Os filmes que apresentam uma menor variabilidade de valores correspondem, quer a programas de fim-de-semana, quer dos dias laborais do final da noite.

Os filmes que apresentam maiores valores são os transmitidos aos Domingos com início entre as 17H-18.15H (14,9%), transmitidos ao Sábado, por volta das 23H (9,4%) e transmitidos ao Sábado por volta das 16.30H-18.30H (7,7%). Neste contexto correspondem essencialmente a programas de fim-de-semana, mas não só e do decorrer da tarde e da noite. O maior valor corresponde ao espaço que teve como contra-programação a transmissão de eventos desportivos por parte do canal estatal RTP1 em determinados dias. Logo, os programas desportivos são os responsáveis pela variabilidade observada em detrimento da eventual irregularidade nos padrões de exposição televisiva.

Em relação aos filmes da TVI também apresentam maiores valores de *TVR_Prog%* e uma maior variabilidade resultante das oscilações horárias e da introdução de programas de maior popularidade. Os filmes da noite para este canal apresentam uma maior estabilidade (<3,2%) em relação aos filmes da tarde. Para a RTP1 os menores intervalos de *TVR_Prog%* estão associados aos filmes de Sábado à tarde (0,3%) e a filmes transmitidos em dias laborais à noite (0,1%) e os maiores valores aos filmes de Domingo transmitidos ao final da tarde (21,8%). Os maiores valores de variabilidade resultam da transmissão de programas de maior popularidade e das oscilações horárias dos programas.

□ Os Programas do Horário Nobre

Em relação aos restantes programas, os maiores valores de variabilidade para o canal SIC e TVI correspondem às telenovelas do horário nobre ou do início da noite. Para o canal estatal RTP1 os maiores valores correspondem aos concursos de pré-*prime-time* e da noite. As faixas horárias a partir das 19H são sem dúvida as que apresentam uma maior variabilidade de valores. Para o canal SIC a maior variabilidade de valores, com um valor aproximado de 10%, corresponde às duas principais telenovelas da noite da semana e à telenovela de pré-*prime-time* (13,6%), seguindo-se a última telenovela da noite. Para o canal TVI os maiores valores correspondem à principal telenovela da noite no decorrer de toda a semana, atingindo para o período em análise o valor de 17,5% (S) e à telenovela *Morangos com Açúcar* (FS) com um valor muito elevado para o período em análise (31%).

Por último, para o canal estatal os valores maximizam-se para o concurso “o Preço Certo” e para o programa de humor de fim-de-semana da noite, seguindo-se os programas equivalentes em dias laborais.

Portanto, pontualmente a variabilidade pode ser muito considerável, especialmente para programas com horários de início mais variáveis do período da noite, horário nobre e pré-*prime-time* e que apresentem como contra-programação determinados jogos de futebol.

Os valores de *TVR_Prog%* em função das durações programáticas parciais podem ser menos variáveis, uma vez que correspondem a intervalos de tempo mais restritos. Do ponto de vista comercial é mais relevante a análise das audiências programáticas parciais, uma vez que os espaços comerciais estão associados a partes de programas e são tarifadas de forma diferenciada, ainda que sejam transmitidos no decorrer do mesmo programa.

Com base na análise dos valores parciais de *TVR_Prog%* verificamos que alguns programas apresentam valores de variabilidade bastante reduzidos, e por vezes a variabilidade global pode ser atribuída apenas a determinadas partes programáticas. Atendendo aos filmes do período da noite é notório o decréscimo verificado no seu decorrer quanto à variabilidade e dimensão de *TVR_Prog%* (P).

Os períodos com maiores níveis de audiências também são os mais variáveis, o que se poderá atribuir em maior grau aos segmentos que apresentam uma menor regularidade na sua exposição televisiva. Quanto aos filmes da tarde após a conclusão do espaço informativo a variabilidade vai decrescendo gradualmente

invertendo-se esta tendência com uma maior proximidade face ao espaço informativo da noite. Os filmes da RTP1 e em especial os que são transmitidos no período da noite podem apresentar uma grande amplitude de valores.

Nas análises de previsão será necessário ter em atenção que as audiências de semana para semana variam de forma diferenciada em função dos programas e da sua localização na grelha. Existem programas que em determinadas faixas horárias e canais apresentam e incutem uma maior variância aos dados. Outros pelo contrário apresentam uma variabilidade muito menor, o que condiciona os instrumentos de previsão mais adequados.

II. Caracterização da evolução da variável *TVR_Período%*

Com base nas análises realizadas identificámos quais são os factores com uma maior repercussão na variabilidade dos dados e a influência decorrente das alterações que se verificam nos horários de transmissão dos programas. Os valores indicam a existência de um reduzido envolvimento com o meio televisivo e com os conteúdos programáticos normalmente transmitidos, para além de uma elevada capacidade do *media* para atrair um significativo número de telespectadores quando os conteúdos transmitidos o justificam. Consequentemente, as motivações dos telespectadores apenas pontualmente aparentam ser instrumentais (*Cooper & Tang, 2009*).

Neste apêndice pretende-se a análise das taxas médias de audiências para o total da TV em função das faixas horárias, considerando intervalos de 30 ou 15 minutos.

Estes períodos correspondem ao horizonte temporal que é utilizado na prática quando se pretende realizar previsões a curto prazo.

Atendendo às semelhanças existentes na evolução desta variável para os vários anos considerados, interessa avaliar até que ponto é possível identificar períodos horários com uma evolução mais comparável. Nos períodos que apresentem uma menor variabilidade da variável em estudo, mais facilmente se poderá prever o seu valor, provavelmente recorrendo a metodologias univariadas de análise de ST. Por outro lado, as faixas horárias onde se verifique uma evolução mais variável devem ser objecto de uma análise posterior mais exhaustiva.

As estratégias de programação e de contra-programação, assim como a composição sociodemográfica da audiência é variável no decorrer do dia. Consequentemente uma vez que está principalmente em causa o desenvolvimento de modelos explicativos de *ratings*, e não univariados, é necessário segmentar a emissão diária em períodos que apresentem tendências similares e que posteriormente devem ser alvo de análises diferenciadas. Devemos verificar se a classificação adoptada tradicionalmente pelo sector quanto à definição das faixas horárias corresponde à que resulta da segmentação dos dados, atendendo à sua variabilidade.

A segmentação dos dados também permite uma representação das séries originais que viabiliza se necessário a posterior aplicação de metodologias do âmbito do “Rule Discovery”, “Novelty Detection” ou do “Motif Discovery” que podem eventualmente contribuir para a previsão deste indicador.

Quanto menor for a variabilidade inter-anual para idênticos períodos horários, mais razoável será podermos concluir que estamos perante um painel que apresenta um consumo essencialmente influenciado por condicionantes temporais e estruturais. Contudo, o comportamento agregado, nem sempre apresenta uma correspondência directa com o que se verifica a nível individual. Se verificarmos a existência de uma considerável estabilidade destes valores, denotando uma eventual maior passividade da audiência, a abordagem univariada para a sua previsão poderá ser equacionada, uma vez que os seus valores serão essencialmente função de factores temporais e em geral esta corresponde à abordagem mais utilizada na prática.

A partir dos dados de *TVR_Prog%* associados a programas procedeu-se ao cálculo de uma estimativa destes valores para períodos horários. Obtemos assim uma série de valores discretos, equitativamente espaçados no tempo. Para a construção das bases de dados considerou-se intervalos de meia hora utilizando valores globais e parciais de *TVR_Prog%*. Consideram-se para além das séries temporais de valores de *TVR%* associados a períodos horários, os intervalos de variabilidade desta variável para cada período e que se traduz em uma base de dados simbólica. As datas em análise coincidem para os vários anos considerados evitando-se assim diferenças que possam ser atribuídas à sazonalidade do consumo. Apesar de se tentar minimizar as desigualdades existentes nas condições subjacentes que possam influenciar o consumo televisivo, estas não podem ser totalmente excluídas.

Meras alterações climáticas poderá ser suficiente para se alterarem os padrões de consumo televisivo entre os vários anos considerados. As alterações que se possam verificar na grelha, ao longo dos anos, também podem influenciar as diferenças identificadas, sobretudo as que decorrem da transmissão de eventos desportivos. Obviamente que dificilmente se mantêm os factores pessoais e individuais que condicionam os padrões de disponibilidade e de predisposição para ver TV.

IIa. A evolução de *TVR_Periodo%*

Observamos um padrão diário de evolução deste indicador, o qual se diferencia entre o fim-de-semana e os dias laborais. As grandes semelhanças existentes na evolução diária desta variável mantêm-se inclusive para os anos mais próximos. Até certo ponto existe alguma estabilidade nos seus valores para idênticos períodos, tornando este indicador independente da programação e a sua variabilidade pode ser atribuída em maior grau a factores de disponibilidade/ vontade do painel. Programas de grande popularidade podem alterar significativamente a evolução desta variável. Nos gráficos que ilustram a evolução média desta variável observamos dois picos de audiências que correspondem aos espaços informativos e ao horizonte temporal denominado de horário nobre. A restante emissão diária subdivide-se em diferentes períodos com tendências ascendentes ou descendentes do tipo polinomial e linear, estando os valores em determinado instante obviamente correlacionados com os adjacentes, em parte também pelos efeitos de herança existentes entre os programas adjacentes.

A análise dos índices de duplicação consta dos capítulos subsequentes, no entanto segundo a bibliografia internacional estes valores não costumam ser muito elevados. Este factor poderá ser uma das causas para a inexistência de um significativo número de períodos onde se possa observar uma maior estabilidade de valores de *ratings*. A associação entre a evolução dos *ratings* e a duplicação primária das audiências consta do *Apêndice F.4*.

Verificamos que atendendo à evolução média de *TVR_Período%* para os dias laborais, assistimos no decorrer da manhã até ao espaço informativo da tarde e a partir das 16.30H a uma tendência ascendente polinomial e após os dois espaços informativos um decréscimo polinomial no início da tarde e linear no decorrer da noite. Apenas entre as 13H e as 14H e entre as 20.30H e as 22.30H existe uma maior estabilidade na dimensão desta variável, correspondendo estes períodos ao espaço informativo da tarde e à segunda parte do espaço informativo da noite e seus programas sucessores. Correspondem às faixas horárias onde após um dia de trabalho as pessoas recorrem à TV por esta ser um veículo de informação, e posteriormente assistem a um reduzido número de programas de ficção. O abandono no decorrer da noite a nível médio processa-se linearmente. A emissão diária a nível médio poderia ser subdividida nos seguintes segmentos horários: desde o início da emissão até 13H, 13H-14H, 14H-16.30H, 16.30H-20H, 20H-22H e 22H até final da emissão. Não corresponde à divisão que *Fernandes (2001)* apresenta como sendo adoptada na indústria televisiva. Ao fim-de-semana e a nível médio a emissão diária também pode ser subdividida em quatro períodos de transição das audiências, com

tendências polinomiais (7.30H-13H; 13.30H-16.30H; 16.30H-20H) e linear (22.30H-FIM). A evolução de *TVR%* em determinados períodos é menos acentuada e o patamar de audiências médio é meia hora mais longo no período da noite (em dimensão ligeiramente inferior aos dos dias laborais, 2-3%), mas idêntico ao semanal à hora do almoço (em dimensão superior ao semanal em 5%). Uma vez mais por volta das 16.30H verifica-se um ponto de inflexão na evolução de *TVR_Período%*. Portanto, sem dúvida que os períodos horários se podem diferenciar quanto às taxas de crescimento e decréscimo, estando os valores desta variável em determinado instante correlacionados com os adjacentes. A proximidade de valores nas tendências médias não é, no entanto, condição suficiente para a variabilidade de valores, ano após ano, ser reduzida para idênticos períodos horários.

IIb. Variabilidade da variável *TVR_Período%*

A maior disponibilidade do público ao fim-de-semana gera uma maior variabilidade na evolução desta variável ao longo de todo o dia até ao período que antecede o espaço informativo da noite, sensivelmente as 19.30H, após o qual existe uma audiência total mais estável. Ao fim-de-semana a existência de uma audiência sem condicionantes laborais reduz a variabilidade observada ao nível dos períodos de maiores audiências no período da noite, mas a partir das 22.30H os valores tendem a equiparar-se, ainda que o período semanal apresente maiores valores médios a partir das 22.30H.

Uma vez que ao longo da semana as pessoas que não exercem uma actividade profissional são as que podem assistir maioritariamente à televisão até às 19H os padrões de consumo semanais apresentam uma maior estabilidade, ou seja, existe uma audiência mais regularmente disponível.

Os valores mínimos semanais de *TVR%_Período* são inferiores aos de fim-de-semana até às 20H, e posteriormente equivalentes ou ligeiramente superiores. Os valores máximos semanais de *TVR_Período%* são inferiores até às 19H e posteriormente inverte-se esta tendência.

Verificamos que durante os dias laborais a variabilidade dos valores é máxima entre as 19H e as 21.30H (13% a 20%) considerando a totalidade da grelha, mas para as restantes faixas horárias, os valores são inferiores aos 6%. Ao fim-de-semana, excluindo o período que decorre entre as 17H e as 20H, os valores de variabilidade ficam aquém dos 10%. Em geral a variabilidade de valores é superior ao fim-de-semana.

Ao fim-de-semana os valores médios, são superiores à hora do almoço, mas inferiores no período de maiores audiências da noite, relativamente aos valores dos dias laborais. Por exemplo, para os anos de maior variabilidade desta variável os valores são no máximo de 15% para os dias laborais e de 10% ao fim-de-semana.

IIC. Evolução Inter-Anual de *TVR_Período%*

Comparando a evolução inter-anual verificamos que o diferencial dos valores de *TVR_Período%* varia em função da distância inter-anual, faixas horárias consideradas e consoante nos referimos a dias laborais ou ao fim-de-semana. As faixas horárias não são equiparáveis quanto à sua propensão para se verificarem alterações nos padrões de disponibilidade do painel.

A hora do almoço apresenta valores de variabilidade quase duplos aos da noite para os dias laborais. O fim-de-semana e a hora de jantar dos dias laborais apresentam uma maior estabilidade de valores entre os vários anos considerados, sendo estes também os horizontes temporais mais relevantes do ponto de vista comercial.

Portanto, verificamos que a nível global existe no decorrer dos anos uma maior estabilidade nos padrões agregados da exposição em períodos de lazer do que em dias laborais. Os períodos que correspondem a uma actividade laboral apresentam um comportamento inter-anual mais diferenciado, apesar da exposição televisiva poder estar associada em maior grau a quem não exerce uma actividade profissional, ou pelo menos por conta de terceiros.

As manhãs diferenciam-se por apresentarem uma variabilidade superior ao fim-de-semana, mas todas as faixas horárias apresentam uma menor variabilidade inter-anual para os períodos de lazer.

Independentemente do ano considerado a evolução diária desta variável é muito semelhante, assim como a repercussão de determinados programas. Por exemplo, atendendo aos dados do ano 2003 a evolução média diária semanal e de fim-de-semana é idêntica à que se calculou para o ano 2004. Se considerarmos o horário compreendido entre as 20H e as 22H a variabilidade de valores entre 2004 e 2003 não excede os 2,2% ao longo de toda a semana, mas considerando o ano de 2002 atinge os 4,2% de variação ao fim-de-semana e de 4,6% em dias laborais e considerando a globalidade dos programas. Verificamos que as oscilações são superiores entre as 20H-20.30H. Considerando a hora do almoço a variabilidade inter-anual é reduzida ao fim-de-semana, mas em dias laborais apresenta valores entre os 6,5% e os 9,6%.

O conteúdo programático poderá em parte influenciar este indicador para determinados períodos horários, mas sem dúvida que o público deve em parte adquirir determinados hábitos de consumo, desenvolvendo uma selecção programática mais rotineira que será tanto mais estável quanto mais o permitirem os factores pessoais subjacentes e imperarem as motivações ritualísticas. Nos capítulos de análise de duplicação das audiências avalia-se a regularidade da exposição. Aparentemente estes valores reflectem essencialmente possíveis alterações nos factores de disponibilidade e vontade de ver TV do painel.

Consequentemente, em períodos de lazer a variabilidade observada é mais reduzida, apesar de nestes períodos se verificarem maiores alterações na grelha de programas e ao nível da contra-programação. Esta tendência é válida para os períodos de maiores audiências. Os maiores valores de variabilidade não estão associados aos períodos de maiores audiências da noite.

A menor variabilidade, em geral associada ao fim-de-semana, poderá permitir-nos concluir que a discrepância de valores observada a nível inter-anual resulta em maior grau de factores de disponibilidade do painel e em menor grau da repercussão do conteúdo programático, desde que excluindo os eventos desportivos.

Quanto mais distanciados estiverem os anos em análise, mais facilmente podemos incorrer em maiores diferenciais de valores, não só pelas alterações verificadas nas grelhas, como também ao nível do padrão individual e global de exposição ao *media*, uma vez que no decorrer dos anos vão existindo alterações na constituição do painel.

Os períodos mais penalizados correspondem à hora do almoço e à tarde durante os dias laborais, contudo a hora do almoço é mais favorecida. No período semanal a nível global existe uma maior divergência entre os vários anos considerados, uma vez que será mais difícil manterem-se os padrões de disponibilidade da audiência, apesar de uma fracção do painel não desenvolver uma actividade laboral.

	Var_Tvr_04/03_S	Var_Tvr_04/03_FS	Var_Tvr_04/02_S	Var_Tvr_04/02_FS
9H	-0,50	7,70	-0,80	6,40
9HM	-2,50	6,00	-4,70	6,55
10H	-0,10	5,18	-2,20	5,08
10HM	-2,50	2,90	-6,50	3,95
11H	-4,30	6,80	-6,20	7,85
11HM	0,25	4,53	-1,55	4,73
12H	-8,90	7,60	-6,50	8,50
12HM	-3,20	1,70	-2,50	-1,30
13H	-8,30	2,70	-6,40	0,90
13HM	-7,05	1,70	-8,20	0,00
14H	-9,50	0,90	-9,60	-2,60
14HM	-7,10	1,78	-8,20	-2,08
15H	-6,40	3,80	-9,70	1,65
15HM	-4,40	0,63	-8,00	-1,33
16H	-5,85	1,40	-10,25	-1,10
16HM	-5,10	2,00	-8,20	-3,20
17H	-3,40	4,40	-6,90	-0,70
17HM	-3,70	8,98	-6,05	5,28
18H	-3,10	7,50	-3,00	1,60
18HM	-4,90	9,10	-2,90	5,20
19H	4,90	4,40	10,00	5,90
19HM	4,90	6,80	2,70	5,90
20H	1,90	1,40	4,60	4,20
20HM	-1,80	-0,80	3,30	1,90
21H	1,40	1,30	2,95	3,35
21HM	-0,30	-2,20	2,90	-1,70
22H	-6,80	-0,50	-1,20	0,30
22HM	-3,10	-1,80	-2,10	-1,10
23H	-4,20	4,80	-2,40	2,00
23HM	-3,00	3,30	-3,90	3,80
24H	-2,30	0,40	-2,10	1,70

	Var_Tvr_04/03_S	Var_Tvr_04/03_FS	Var_Tvr_04/02_S	Var_Tvr_04/02_FS
9H	-0,50	0,90	-0,80	-0,40
9HM	-2,50	0,00	-1,30	0,55
10H	-0,10	0,68	-0,90	0,58
10HM	-2,50	-1,30	-4,00	-0,25
11H	-4,30	-0,30	-3,70	0,75
11HM	0,25	-0,28	-1,55	-0,08
12H	-8,90	1,70	-6,50	2,60
12HM	-3,20	1,40	-2,50	1,80
13H	-8,30	2,70	-6,40	0,90
13HM	-7,05	1,70	-8,20	0,00
14H	-9,50	0,90	-9,60	-1,80
14HM	-7,10	1,78	-8,20	1,03
15H	-6,40	3,80	-9,70	1,65
15HM	-4,40	0,63	-8,00	-1,33
16H	-5,85	1,40	-10,25	-1,10
16HM	-5,10	2,00	-8,20	1,40
17H	-3,40	4,40	-6,90	4,90
17HM	-3,70	8,98	-6,05	10,08
18H	-3,10	7,50	-3,00	7,70
18HM	-4,90	9,10	-2,90	11,30
19H	4,90	9,00	10,00	10,30
19HM	4,90	9,20	2,70	9,60
20H	1,90	1,40	4,60	6,60
20HM	-1,80	-0,80	3,30	1,90
21H	1,40	1,30	2,95	3,35
21HM	-0,30	-2,20	2,90	-1,70
22H	-6,80	-0,50	-1,20	0,30
22HM	-3,10	-1,80	-2,10	-1,10
23H	-4,20	4,80	-2,40	2,00
23HM	-3,20	3,30	-4,10	3,80
24H	-2,30	0,40	-2,10	1,70

Tabela F.1.1.5.II.1 e Tabela F.1.1.5.II.2 – Variabilidade de valores da variável TVR_Periodo%

Ao fim-de-semana mais facilmente se pode verificar uma maior estabilidade e o desenvolvimento de rotinas por inexistência de condicionantes laborais.

Os conteúdos ao longo dos anos não aparentam ter uma grande repercussão nos valores médios deste indicador desde que não se considerem programas pontualmente muito atractivos, o que corrobora o pressuposto de alguma predominância das motivações ritualísticas. No entanto, se considerarmos a totalidade da amostra mantêm-se os resultados, a nível médio. A transmissão de programas de maior popularidade, a nível médio, não influencia significativamente os resultados. Durante os dias laborais a introdução destes programas na grelha altera a sua evolução média para os correspondentes períodos horários no máximo em 0,25% e ao fim-de-semana em uma unidade percentual.

No entanto, para previsões concretas e diárias a influência destes programas deve ser considerada. Atendendo aos resultados, os valores dos *ratings* parecem ser essencialmente função dos períodos horários a que estão associados e da capacidade dos canais em conseguirem maximizar os seus *shares* de audiências face ao volume global de exposições. Alterações nos horários de transmissão, obviamente, vão alterar os respectivos índices de audiências. Quanto mais notórios forem os comportamentos de lealdade ao canal e programáticos menor será a variabilidade observada nos *ratings* face a alterações horárias ou como consequência da contra-programação.

Consequentemente, é necessário a segmentação da emissão diária, segundo a evolução verificada neste indicador, reduzindo o carácter temporal dos dados a uma categorização nominal uma vez que está em causa o desenvolvimento de modelos explicativos de audiências.

A introdução de programas de grande popularidade, nomeadamente determinados jogos de futebol transmitidos essencialmente por parte do canal RTP1 podem alterar os valores desta variável para períodos horários equivalentes. Estes programas podem contribuir para o acréscimo do valor desta variável associado aos programas em contra-programação e no caso dos programas adjacentes poderá ter ambos os efeitos. Quer dizer, se os programas adjacentes forem antecipados em princípio saem penalizados, atendendo à selecção programática mais selectiva que se verifica para o canal estatal. Se for adiado o início dos programas adjacentes, estes podem ser beneficiados em virtude da existência de efeitos de herança entre os programas adjacentes. No entanto, a dimensão dos efeitos de herança é função dos novos suportes horários que se associam a determinado programa e também da contra-programação. A repercussão destes programas também depende dos períodos considerados, sendo mais notória em períodos de maiores audiências. Porém estes últimos também correspondem aos períodos seleccionados pelas estações televisivas para a transmissão dos mesmos com o intuito de maximizarem os respectivos *shares* de audiências.

Portanto, desde que se verifiquem significativas alterações na grelha de programas existe alguma estabilidade nos valores desta variável a nível médio para períodos horários equivalentes. Em princípio será possível recorrer a modelos univariados de previsão deste indicador.

III. A Influência das Variáveis Sociodemográficas na evolução de *TVR_Período%*

Não existe consenso relativamente a se as características sociodemográficas influenciam significativamente os padrões diários e semanais de exposição televisiva e a predisposição para se desenvolverem rotinas de consumo televisivo, influenciando o grau de estabilidade inerente à exposição televisiva e a *TVR%*.

Alguns autores defendem que estas variáveis apresentam uma ligeira repercussão nos factores de disponibilidade, o que não influencia obrigatoriamente de forma significativa a exposição televisiva (Cooper & Tang, 2009). Do ponto de vista da previsão, uma vez que as estações televisivas, por vezes realizam estas análises por público-objectivo, será necessário diferenciar os segmentos atendendo à sua exposição televisiva.

Consequentemente, as análises a realizar apresentam os seguintes objectivos:

➤ Diferenciar os segmentos sociodemográficos representados nas bases de dados de audimetria relativamente ao seu grau de exposição ao meio televisivo, identificando quais são os segmentos que apresentam evoluções mais semelhantes à evolução da totalidade do painel e entre si.

As análises são realizadas para a globalidade da emissão diária, assim como para períodos horários mais restritos. Calcularam-se as distâncias euclidianas entre as ST originais associadas a valores programáticos e os coeficientes de correlação entre as mesmas, permitindo estes últimos realizar o *clustering* das ST recorrendo a algoritmos hierárquicos e ao algoritmo K Média. O número de *clusters* foi definido com base na análise em Componentes Principais.

Posteriormente, definindo como objectos simbólicos os períodos horários com uma duração de 30' e caracterizados por variáveis da tipologia intervalo $([\alpha, \beta])$ que representam a amplitude de valores de *TVR_Período%* para os vários segmentos representados nas bases de dados aplicaram-se medidas de dissimilaridade e algoritmos de *clustering* para objectos simbólicos booleanos. Verificamos, assim quais são os segmentos mais semelhantes em função da variabilidade de valores de *TVR_Período%* e considerando a totalidade dos dados.

Segundo as análises simbólicas, os segmentos que a nível global apresentam valores mais diferenciados relativamente ao universo televisivo são a região do interior, a classe social D, o sexo feminino, as faixas etárias a partir dos 55 anos e os lares sem TV Cabo e com menores valores, a classe AB e os Lares com TV cabo. Facilmente podemos estabelecer uma associação entre as classes sociais mais altas e um maior grau de adesão à TV Cabo. Logo, o acesso a uma maior diversidade de canais e o facto de a TV deixar de ser um serviço gratuito não implica obrigatoriamente um maior exposição televisiva.

Na literatura da especialidade, por vezes, é mencionado que estando um custo associado ao meio televisivo poderia implicar um maior grau de exposição. No entanto, este segmento apresenta a nível nacional o menor nível de exposição televisiva. Uma vez que poderá não se identificar com os conteúdos que constam da grelha de programação generalista poderá recorrer a canais alternativos que permitem satisfazer outro tipo de necessidades e interesses. A exposição ainda que menor leva a uma selecção mais selectiva dos conteúdos. A classe designada de A inclui os quadros médios e superiores com licenciaturas, mestrados e doutoramentos. A classe designada de B abrange os técnicos especializados e os pequenos proprietários, assim como os empregados dos serviços, comércio e administração e os trabalhadores qualificados e especializados que regra geral apresentam no mínimo cursos profissionais.

Com uma maior exposição televisiva consta o segmento feminino e os lares com menores rendimentos, e consequentemente em geral sem TV Cabo. A classe designada de D inclui os indivíduos com até ao 9^a ano de escolaridade, os que são não activos, as domésticas e os trabalhadores não qualificados e não especializados.

As regiões do interior do país também são as que mais se diferenciam da evolução dos indicadores de audimetria para o universo televisivo. Inclui cidades dos Distritos de Aveiro, Braga, Bragança, Castelo Branco, Coimbra, Guarda, Leiria, Porto, Santarém, Viana do Castelo, Vila Real e Viseu.

Atendendo, às classes sociais, ao fim-de-semana, a classe D diferencia-se no decorrer da totalidade da emissão diária face aos restantes segmentos com maiores valores de exposição televisiva e em dias laborais, essencialmente, no decorrer do horário nobre e em menor grau à hora do almoço. A classe AB também se diferencia com menores valores de exposição nos períodos de maiores audiências e apresenta um comportamento mais homogéneo no decorrer da semana ao longo do período em análise. Os quadros médios e superiores com maiores níveis de instrução apresentam uma maior uniformidade de exposição televisiva no decorrer da semana, não se verificando um acréscimo significativo ao fim-de-semana. A classe D, por sua vez, ao fim-de-semana aumenta significativamente a sua exposição televisiva no decorrer de todo o dia, diferenciando-se também em dias laborais após as 18H.

O segmento masculino em dias laborais minimiza os valores da exposição televisiva face ao segmento feminino essencialmente nos períodos de maiores audiências nomeadamente no decorrer do horário nobre. Ao fim-de-semana tende a equiparar-se ao segmento feminino em grande parte da emissão diária. No decorrer da semana apresenta valores de exposição coincidentes com os do sexo feminino a partir das 24H, e por vezes ainda se verifica uma inversão apresentando valores superiores. No final da noite quem assiste em maior grau à TV corresponde aos segmentos D e C1 e, logo, talvez devido ao denominado “co-viewing”. Existindo em muitos lares apenas um televisor, a tendência será para assistir em conjunto aos programas, apresentando valores equiparáveis de *TVR_Período%*.

Em dias laborais existe uma maior proximidade entre a evolução de TVR% para as faixas etárias até aos 25 anos que minimizam estes valores e um segundo conjunto que os maximiza e que inclui as faixas etárias superiores aos 55 anos.

As faixas etárias dos 35-55 anos ocupam um lugar intermédio. As maiores faixas etárias, em geral, recorrem em maior grau a este *media*, independentemente das motivações subjacentes, que podem ser a procura de informação ou de entretenimento, e por vezes inclusive como uma forma de fuga à solidão. Do ponto de vista estatístico podemos agregar as faixas etárias dos 55-65 anos e o público que apresenta mais de 65 anos. Os mais jovens até aos 25 anos apresentam o menor grau de exposição televisiva, uma vez que devem optar por outras actividades lúdicas no lar ou no exterior. Uma vez que o painel apresenta em menor grau acesso à TV_Cabo o grau de penetração de outras tecnologias de informação, como, por exemplo, a Internet que poderiam exercer uma maior atractividade nas faixas etárias mais jovens também deverá ser diminuto. A partir do 25 anos verifica-se um acréscimo ao nível da exposição televisiva.

Ao fim-de-semana as duas maiores faixas etárias, a partir dos 55 anos diferenciam-se nos períodos de maiores audiências em vez de no decorrer da totalidade da emissão diária e os mais jovens até aos 14 anos também apresentam valores mais diferenciáveis da média observada no decorrer da manhã, uma vez que a grelha de programação também se destina em maior grau a estes segmentos.

Em dias laborais os lares com TV_Cabo apresentam menores valores para os períodos de maiores audiências, sendo esta diferença menos acentuada ao fim-de-semana.

Para as análises de *TVR_Período%* associadas a segmentos uma vez que estas apresentam uma evolução diária que segue uma evolução sinusoidal também se avaliou a possibilidade de transformação das séries temporais para o domínio da frequência, utilizando as metodologias de DFT e o DWT por forma a facilitar a sua comparação e a avaliação do grau de semelhança. No entanto, estas metodologias não acrescentaram nenhuma mais valia às análises realizadas, apresentando as metodologias simbólicas resultados mais facilmente interpretáveis.

➤ Verificar quais são os segmentos que apresentam uma maior consistência nos seus perfil de exposição televisiva.

Para além de identificar padrões médios de exposição intrínsecos a determinados segmentos, será relevante identificar o grau de estabilidade dos mesmos no decorrer da semana para os diferentes segmentos. Recorrendo ao algoritmo K_Média e à Análise em Componentes Principais identificaram-se segmentos com evoluções diferenciadas. Estas análises foram realizadas em função do dia de semana. A identificação dos agrupamentos com um grau de exposição reduzido, médio e elevado consta do *Quadro F.1.1.5.III.1*. Estas análises foram realizadas aos indicadores agregados.

Dia de Semana I			
Segmentos	Faixas Horárias		
	8H-8.30	..	24H-24.30H
Segmento 1	Valor	..	Valor
	Tvr%		Tvr%
..	..	..	..
..	..	..	..
Segmento N	..	..	..
Dia	Segmento		
	Seg_1	...	Seg_M
Segunda	A	...	M
Terça	M	...	M
Quarta	A	...	M
Quinta	A	...	A
Sexta	A	...	A
Sábado	B	...	B
Domingo	M	...	M
	Classe AB	...	Feminino
Sexta 1	A	...	A
...	...	...	...
Sexta N	M	...	A
Classifica- ção Final	A que Maximiza a Frequência	...	A que Maximiza a Frequência

Quadro F.1.1.5.III.1 – Identificação de segmentos mais coincidentes ao nível da sua exposição televisiva

Por exemplo, para cada uma das sextas feiras que integram a base de dados foi atribuído a cada segmento o *cluster* R, M, A, representando uma exposição reduzida, média ou alta, sendo a classificação final função dos *clusters* que maximizam as frequências de ocorrência para cada um dos segmentos (Quadro F.1.1.5.III.1)

Por exemplo, às sextas associamos uma exposição mais elevada apenas a quatro segmentos: à classe D, faixas etárias 55-64 e +64 e às Donas de Casa e com uma exposição reduzida constam as classes sociais AB e C1, sexo masculino, faixas etárias até aos 46 anos, lares com TV Cabo e Não Dona de Casa.

Os restantes segmentos apresentam uma exposição intermédia.

Posteriormente, pretendeu-se verificar se existe alguma estabilidade no decorrer da semana para os diversos segmentos. A cada segmento e dia atribui-se a classificação que maximiza as frequências calculadas. Verificamos que no decorrer dos dias laborais a atribuição de segmentos a *clusters* apresenta uma forte analogia, levando a pressupor a existência de padrões de consumo que são maximizados pela classe social mais baixa e por parte das maiores faixas etárias a partir dos 55 anos.

Os padrões de disponibilidade acabam por definir os padrões de exposição televisiva, diferenciando-se os três níveis de exposição apenas relativamente à magnitude e em menor grau quanto à forma de evolução deste indicador.

Ao fim-de-semana os padrões de consumo para os diversos segmentos apresentam alguma divergência.

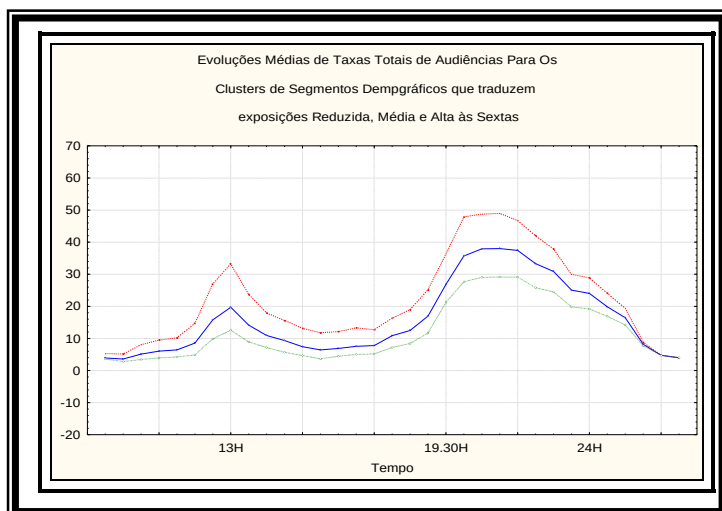


Gráfico F.1.1.5.III.1 – Evoluções médias de *Tvr_Periodo%* para os segmentos que apresentam uma maior ou menor exposição televisiva

IV. Influência dos Espaços Comerciais na evolução de *Tvr_Periodo%*

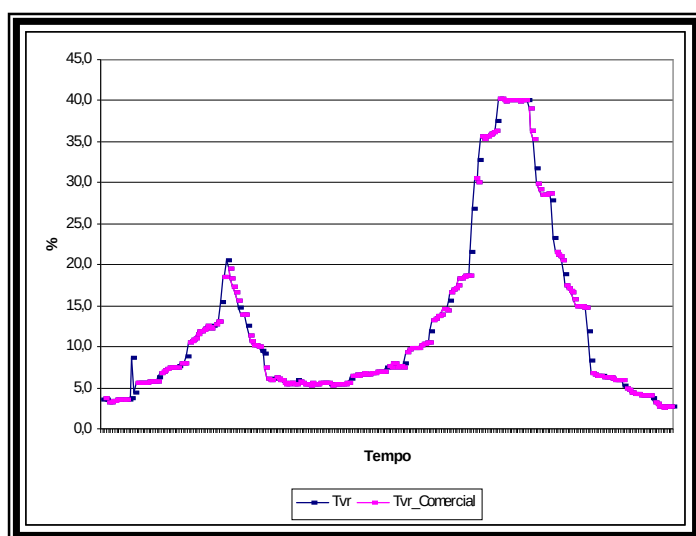


Gráfico F.1.1.5.IV.1 – Evolução de *Tvr_Periodo%* no decorrer dos espaços comerciais

Também, é relevante verificar em que medida os espaços comerciais podem levar a uma alteração na evolução desta variável. Como se pode observar no Gráfico F.1.1.5.IV.1 em função das faixas horárias a relação existente entre os valores de *Tvr_Prog%* e os valores comerciais diferencia-se.

A evolução deste indicador no decorrer dos espaços comerciais segue a evolução diária esperada, apresentando uma tendência ascendente no decorrer da manhã e após as 16H e uma evolução descendente após o Jornal da Tarde e no horário nobre. Coloca-se a dúvida de qual será o grau de repercussão do planeamento comercial quanto ao ritmo de redução de *TVR%*. Verificamos que no período da noite existe uma rápida redução deste indicador. Espera-se a existência de alguma dependência dos valores desta variável quando associada a intervalos face à envolvente programática, assim como, se verifica ao nível dos *ratings*.

No Apêndice F.7 realizam-se algumas análises no âmbito comercial. A previsão das audiências comerciais é fundamental uma vez que com base nestas estimativas são decididos os preços e avalia-se o grau de competitividade dos espaços comerciais face aos da concorrência. Os anunciantes e as centrais de meios precisam conhecer estas previsões quando planificam as suas campanhas publicitárias para estarem aptos a seleccionar os tempos mais rentáveis. As previsões têm que ser realizadas para a totalidade dos canais e num só dia, já que servem de suporte à selecção das combinações óptimas de suportes. Uma vez que os espaços mais rentáveis vendem-se rapidamente estes sistemas devem ser rápidos, o que condiciona fortemente as metodologias a serem utilizadas.

APÊNDICE F.1.2. ANÁLISES EXPLORATÓRIAS ÀS TAXAS MÉDIAS DE AUDIÊNCIAS E AOS VALORES DE *SHARES* DE AUDIÊNCIAS

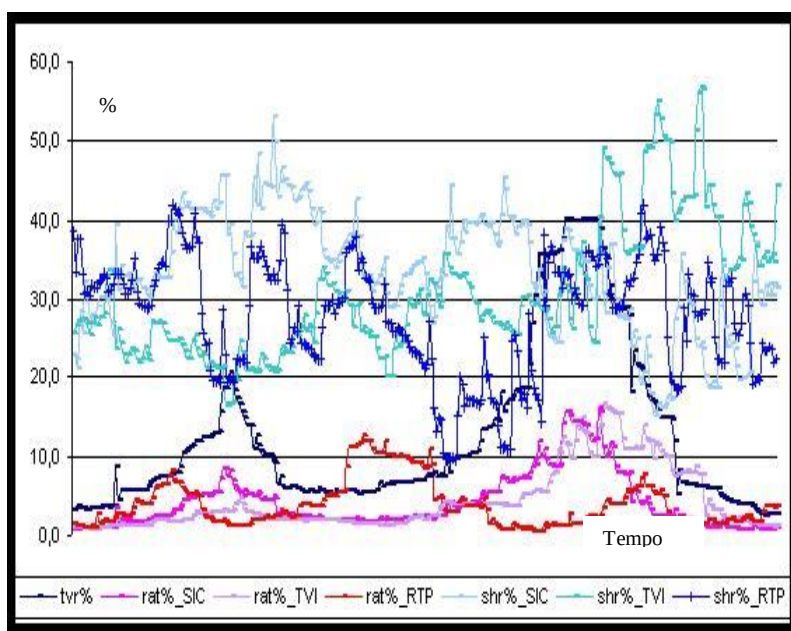
F.1.2.1. Introdução

O índice de audiência mais relevante para o planeamento de *media* e para a estruturação da grelha de programas corresponde às taxas médias de audiências, também designadas por *ratings*. Este indicador pondera os telespectadores em função da duração da exposição ao evento considerado e encontra-se dependente de inúmeros factores, entre os quais os factores estruturais que condicionam a exposição ao *media* e a interacção com os suportes publicitários, por exemplo, as características da programação e de estruturação da grelha de programas e os comportamentos estruturais.

Do ponto de vista das bases de dados de audimetria, as taxas médias de audiências resultam do produto das taxas médias de audiências para o total da televisão e de valores de *shares*. A exposição do painel ao *media* influencia este indicador, assim como os valores de *shares*, que por sua vez dependem da capacidade dos canais em conseguir favorecer os efeitos de herança entre programas e os comportamentos de lealdade programática e ao canal atenuando a repercussão da contra-programação e de outras estratégias de programação face ao horizonte temporal em que os telespectadores apresentam disponibilidade para assistir ao *media*.

A cobertura total ou "reach" representa o conjunto de pessoas que viram pelo menos um segundo o suporte em análise, ou seja, uma emissão ou período de tempo. A audiência média ou *rating*, representa o número de pessoas presentes, em média, durante a emissão ou período de tempo e traduz-se na média de todas as audiências instantâneas. O *share* de audiência, corresponde à relação observada entre a audiência média de uma emissão ou de um horário face à audiência média da televisão durante o mesmo período de tempo, sendo designado neste último contexto por taxas médias de audiências para o total da televisão.

Consequentemente, a audiência de uma emissão não é medida por apenas um número. Muitas vezes, reduz-se a noção de audiência a um só indicador, nomeadamente à audiência média. A relação entre os vários indicadores será avaliada no *Apêndice F.6*.



Os valores *de ratings*, *shares* e de *Tvr%* encontram-se inter-relacionados e por si só dependentes de inúmeros factores, nem sempre quantificáveis, previsíveis ou conhecidos. Da sua interacção que será variável no decorrer do tempo depende o desempenho da grelha de cada um dos canais.

Gráfico F.1.2.1.1 – Evolução dos indicadores *Tvr%*, *Rat%* e *Shr%* considerando o universo televisivo

Apesar do investimento em publicidade televisiva se encontrar fortemente dependente da previsão deste indicador, outras perspectivas de análise também são viáveis, uma vez que se pretende maximizar a exposição do público-objectivo seleccionado ao plano que integra um conjunto de suportes publicitários. A selecção dos suportes poderá ser realizada através da identificação dos conteúdos que no espaço factorial decorrente da realização de análises de Correspondência ou análises MDS se aproximam em maior grau do público-alvo da campanha publicitária.

Do ponto de vista da previsão dos *ratings* não podemos *a priori* assumir como determinante a necessidade de constituir os restantes indicadores de audimetria como variáveis explicativas e com um simultâneo potencial predictivo, o que também dependerá dos períodos horários e dos dias considerados nas análises e do desempenho e características da programação envolvida.

Nem sempre deverá ser obrigatório a desagregação das análises ao nível dos dados de audimetria para segmentos ou a nível individual com o objectivo de conseguir prever este indicador para o universo televisivo que corresponde à variável mais frequentemente utilizada no planeamento de *media*.

Por exemplo, de acordo com o *Gráfico F.1.2.1.1*, verificamos que os valores de *shares* podem apresentar uma forte variabilidade no decorrer do dia, diferindo substancialmente entre os diferentes canais e contrastando com a evolução verificada ao nível dos *ratings*, que por vezes se aproxima da evolução das taxas médias para o total da TV e que pode apresentar uma maior estabilidade. Os *ratings* comparativamente às taxas médias de audiências para o total da TV devem estar mais dependentes da estruturação da grelha e da interacção do painel com os conteúdos transmitidos. Consequentemente, apesar da variabilidade observada nos valores de *ratings* poder transcender a verificada para a variável *Tvr%*, talvez a previsão deste indicador no decorrer de algumas faixas horárias possa ser realizada recorrendo a análises univariadas ou estabelecendo uma correspondência linear ou não linear apenas com a evolução de *Tvr%*. No *Apêndice F.2* avaliam-se estas alternativas de análise e neste apêndice consta a caracterização da evolução dos *ratings* e dos *shares* de audiências.

Se o objectivo das análises se traduzir na previsão deste indicador será necessário realizar análises exploratórias com o intuito de caracterizar a evolução dos *ratings* e de valores de *shares*, avaliando o seu grau de variabilidade para um período característico a nível médio do consumo televisivo. Paralelamente a identificação dos factores responsáveis por uma maior variabilidade de valores, também, é relevante. Em função dos resultados será então viável identificar os contextos em que a previsão de valores de *ratings* implica a necessidade de serem considerados os dados para segmentos de telespectadores ou as interacções existentes entre as diferentes variáveis de audimetria e que neste último contexto constam posteriormente do *Apêndice F.6*. Uma vez mais foi muito difícil a identificação de trabalhos publicados neste âmbito ainda que internacionais.

Uma vez que a previsão de valores de *ratings* pode estar dependente do cálculo dos valores de *shares*, neste apêndice constam algumas conclusões relativamente à evolução dos *shares* de audiências. As análises quantitativas de análise de *shares* encontram-se subvalorizadas, apesar de contribuírem para que melhores resultados sejam alcançados face a uma mera utilização de consultas a painéis de opinião. Apenas foram identificadas duas publicações referentes à análise quantitativa dos valores de *shares* de audiências (*Gensh & Shaman, 1980; Henry & Rinne, 1984*) que promovem a sua previsão e que no segundo caso avaliam simultaneamente a repercussão de algumas variáveis na sua evolução. *Gensh e Shaman (1980)* propõem quatro abordagens que se baseiam apenas em médias com os necessários ajustamentos dos seus valores históricos. Os autores não realizam qualquer análise para promover a identificação de variáveis de programação e controláveis que possam explicar a evolução dos *ratings* de cada um dos programas que integram a amostra. Em oposição *Henry e Rinne (1984)* recorrem às análises Logit para tentar prever os valores *shares* de audiências por utilização de algumas variáveis explicativas de audimetria. De acordo com *Danaher et al. (2011)* e *Gensh e Shaman (1980)* são escassos os trabalhos existentes nesta área. Por exemplo, *Gensh e Shaman* referem que:

Unfortunately, the research performed by these organizations is usually viewed as proprietary and there is little information available in the public domain on predictions of television ratings. In discussions with a number of practitioners, we have noted that the proprietary research seems to focus heavily on program content. ... To best of our Knowledge there are no previous published studies that attempt to analyse television-viewing patterns over time. ... It is our contention that bench-mark estimates derived from time series models can provide a framework within which expert opinion on program content and life cycles can be incorporated, to provide predictions superior to those base solely on expert opinion.

Por conseguinte, antes do desenvolvimento dos modelos explicativos de audiências também é relevante a caracterização da evolução dos valores dos *shares* de audiência e de *ratings*, realizando análises univariadas. Será necessário avaliar o grau de independência desta variável face à programação; qual a repercussão das variáveis temporais e qual a influência decorrente da introdução de programas que apresentam maiores valores de exposição na sua evolução. Recorrendo a análises univariadas aplicam-se metodologias de diversos âmbitos com o intuito de tentar realizar a sua caracterização.

De acordo com o seguinte parágrafo (*Gensch & Shaman, 1980*), as análises MDS também podem contribuir para a previsão deste indicador e no *Apêndice F.1.2* caracterizam-se os valores de *ratings* para os diversos segmentos, por forma a avaliar se estes discriminam a evolução observada nos valores de *ratings*.

The Gensch & Ranganathan study has shown that while segments of the population, defined in terms of demographics or past purchasing histories, can have statistically significant mean differences in their preferences for the various types of television programs, there is usually considerable heterogeneity within a segment. ... Through the use of such variables as age, sex, occupation, participation in community activities, ... one may identify broad market segments with different seasonal and weekly viewing patterns. ... Prior published work on television viewing uses static models such as factor analysis and multidimensional scaling, which also attempts to analyze and predict television ratings in terms of program content.

F.1.2.2. Objectivos

A presente tese visa essencialmente o desenvolvimento de análises explicativas de valores de *ratings* programáticos e comerciais, que permitam a identificação dos factores que os condicionam e das interacções existentes e que possivelmente viabilizem a sua posterior previsão. Neste apêndice pretende-se nomeadamente realizar as seguintes análises:

(1) Caracterizar a evolução e grau de variabilidade de *Rat_Prog%*, ou seja, dos valores das taxas médias de audiências quando estão associadas à duração de um programa. Nomeadamente, verificar qual é a repercussão ao nível dos *ratings* das alterações horárias observadas em determinados programas ou ao nível da contra-programação em função dos períodos horários e canais considerados.

(2) Caracterizar a evolução e avaliar o grau de variabilidade inter-anual e intra-anual observado nos valores de *ratings* associados aos diversos canais, recorrendo a algoritmos simbólicos de medidas de dissimilaridade e de *clustering* de séries temporais. O grau de estabilidade poderá ser função de inúmeros factores, como, por exemplo, dos períodos horários, canais e tipologias programáticas ou popularidade dos mesmos, sendo por conseguinte determinante identificar quais são os factores mais relevantes.

Podemos salientar os seguintes factores com influência na evolução dos indicadores de audimetria: O mês em que o programa é transmitido, uma vez que existe uma forte sazonalidade na exposição televisiva, sendo esta superior nos meses de Inverno; o dia de semana e a hora, o canal, o programa, a contra-programação, a popularidade do programa anterior e do programa sucessor. As análises de previsão devem ter em atenção alguns destes factores com o seu peso correspondente. Cada factor acarreta erros de amostragem e determinada variabilidade, logo quando os resultados são muito ambiciosos acabamos por considerar vários erros com as suas correspondentes ponderações, levando a uma reduzida precisão da previsão. Consequentemente, devem ser seleccionados apenas os factores mais relevantes e prescindir dos que se considerem como sendo acessórios. Nesse sentido, as análises a realizar pretendem identificar os factores que são responsáveis pela maior variabilidade observada.

Algumas das questões em análise corresponde às seguintes:

- Qual a evolução da variável em estudo para cada um dos canais? Qual o seu padrão de variabilidade em função das faixas horárias, dias, anos e canais? Os canais privados face aos canais estatais apresentam uma maior ou menor estabilidade de valores no decorrer dos anos ou a nível mensal? Em que medida o desempenho da programação poderá condicionar os resultados? Que conclusões com base na variabilidade observada ao nível dos *ratings* podem ser retiradas para cada um dos canais quanto a prevalecerem as motivações ritualísticas ou instrumentais no painel de audimetria? Qual o grau de variabilidade nos *ratings* que resulta das oscilações horárias dos programas? Qual a influência dos eventos desportivos no canal de transmissão e nos canais concorrentes e na programação adjacente? Que cenários de evolução dos principais indicadores de audimetria podem surgir nestes contextos?
- Qual a repercussão das variáveis sociodemográficas na sua evolução e no seu grau de variabilidade? Correspondem a variáveis discriminantes da evolução dos *ratings*? Quais os segmentos que apresentam uma maior semelhança relativamente à evolução dos *ratings* e face ao universo televisivo? Quais os segmentos que apresentam uma maior variabilidade de valores de *ratings*? Em que medida o canal e as faixas horárias influenciam os resultados?

(3) Análises de caracterização da evolução da variável *share* de audiências. Os valores de *shares* de audiência de determinado programa podem influenciar os respectivos valores de *ratings*, no entanto as alterações no valor das taxas médias de audiência para o total da TV podem, por vezes, alterar a redistribuição dos valores dos *shares* e apresentar uma menor influência na evolução dos *ratings*. Pretende-se avaliar o grau de variabilidade associado a esta variável e confirmar se se justifica a sua constituição como um factor explicativo da evolução dos dados em análises de previsão de *ratings*.

Atendendo às análises gráficas realizadas, este indicador apresenta uma forte variabilidade e que em geral não é acompanhada pela que caracteriza a evolução dos *ratings*. O diferencial de valores entre os valores de *ratings* e das taxas médias de audiências para o total da TV representa para o período em análise entre 50%-90% dos valores de *Tvr%*. Contudo, referem-se as abordagens descritas em *Gensh e Shaman (1980)* e em *Henry e Rinne (1984)* com as necessárias adaptações.

➤ Qual a evolução e variabilidade da variável *share* de audiência para os canais generalistas? Justifica-se recorrer a este indicador quando se pretende prever as taxas médias de audiências? Ou será preferível realizar análises apenas em função de *Tvr_Periodo%*?

F.1.2.3. Base de dados

A base de dados corresponde a variáveis de caracterização da programação e aos valores de *ratings* e de *shares* programáticos que constam da base de dados Telereport. Consideram-se os valores globais e parciais, consoante se referem à globalidade do programa ou apenas a cada uma das partes que o integram. A taxa média de audiência ou *rating* representa a audiência média por segundo. Para o seu cálculo, cada indivíduo é ponderado relativamente ao tempo de contacto com o evento ou canal, representando segundo a bibliografia da especialidade a percentagem do painel que esteve em média a assistir a determinado evento.

$$\text{Rat\%} = \left(\frac{\text{Total de segundos contactados}}{(\text{Duração do programa} * \text{n.º de indivíduos do alvo})} \right) * 100$$

Situação I					
T1	T2	T3	T4	T5	T6
1		1		1	1
	1		1	1	
1		1		1	1
Situação II					
T1	T2	T3	T4	T5	T6
1				1	1
	1			1	
1				1	1

Contudo, facilmente verificamos que com uma alteração na duração da exposição ao evento, ainda que o número de telespectadores se possa manter, os valores dos *ratings* não seriam equivalentes (Quadro F.1.2.3.1). Na situação I, o total de segundos contactados corresponde a 11 segundos e na situação II, tendo em atenção também três telespectadores, este valor é de apenas 8 segundos. Logo, produziria valores diferentes de *ratings*, apesar de ser idêntico o número de telespectadores.

Quadro F.1.2.3.1 – Cenários de exposição televisiva e cálculo dos valores de *ratings*

A utilização do *rating* como um indicador representativo do valor médio da exposição televisiva do painel de audimetria levanta algumas objecções, porém corresponde à variável mais relevante para o planeamento de *media*. As taxas de cobertura tentam quantificar a exposição considerando a percentagem de indivíduos que viram pelo menos 1 segundo do evento, logo ainda que um indivíduo contacte duas ou mais vezes com um evento, apenas será considerado como um contacto. Segundo *Souchon (2003)*, quase sempre, o número de telespectadores que realmente veem um fragmento de uma determinada emissão é maior que o valor expresso pela sua audiência média, pois muitos telespectadores entram e saem da emissão, apresentando uma exposição com uma reduzida duração. Essa grandeza encontra-se quantificada na taxa de cobertura de um evento.

Ao contrário do que *Fernandes (2001)* salienta, ainda que o número de pessoas fosse idêntico entre dois programas ou períodos, não seriam obrigatoriamente equivalentes os respectivos valores de *ratings*: “... o facto das audiências de programas distintos serem similares, não significa que são os mesmos telespectadores. Pelo contrário, as audiências de dois programas compostos por pessoas diferentes.” Isto não representa uma suposta inconsistência dos dados e apenas com a análise dos ficheiros base de audimetria ou com os indicadores acumulados se poderá esclarecer esta questão.

Foram realizadas algumas análises ao nível da duplicação das audiências e relacionou-se a evolução dos indicadores acumulados de audimetria com a evolução dos *ratings*. Como já foi demonstrado, não existe uma associação entre o número de pessoas expostas a dois programas e a estabilidade dos valores dos seus *ratings*. Quer dizer, quando estão presentes o mesmo número de pessoas, os suportes podem apresentar valores diferentes de *ratings* e um número diferenciado de pessoas pode produzir um valor semelhante de *ratings*, ou não, tudo dependerá da duração da exposição seja esta contínua, ou não. Com base na análise dos índices agregados de duplicação das audiências podemos saber qual é o número de telespectadores partilhado entre dois programas ou eventos. O *share* de audiência de um evento ou canal é calculado a partir do tempo total despendido a ver esse evento ou canal relativamente ao tempo total despendido a ver televisão (canal base). Considera-se nas análises a realizar os valores para o universo televisivo.

$$\text{Shr}\% = \left(\frac{\text{Rat}\% \text{ do evento / canal}}{\text{Tvr}\%} \right) * 100$$

Estas variáveis podem-se referir a valores programáticos globais ou parciais se considerarmos cada uma das partes que integram um programa, assim como ao canal. Quando está associada ao canal, a variável será designada por *Rat_Período%* ou *Shr_Período%* e considera-se como horizonte temporal um intervalo de tempo de 30 minutos de duração.

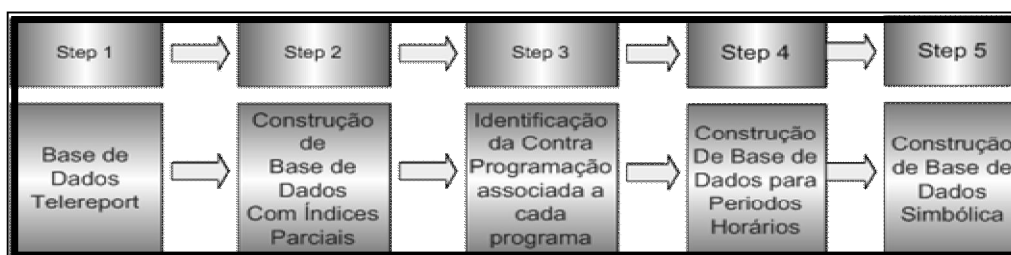


Fig. F.1.2.3.1 – Estruturação de base de dados de *Ratings*

Inicialmente construiu-se a base de dados de índices globais e na sua sequência a base de dados de índices parciais (Fig. F.1.2.3.1). Através do *software* Telereport não é possível construir de forma directa a base de dados de *ratings* parciais, uma vez que os programas globais com apenas uma parte programática são representados apenas com o nível 1 programático, e, logo, são excluídos da amostra dos dados parciais.

Por conseguinte, considerando a base de dados obtida por selecção do nível dois e três programático foi necessário adicionar-lhe todos os programas que apresentam uma única parte. Posteriormente, também foi necessário a correcção das variáveis de caracterização da programação, como, por exemplo, em programas que apresentam várias partes proceder à sua identificação e diferenciar os programas com a mesma designação que são transmitidos em diferentes faixas horárias. A identificação da contra-programação tem que ser realizada programa a programa e recorrendo ao *software* Telereport. Para cada um dos programas seleccionados e discriminados temporalmente ao segundo, surge a identificação do horário de início de cada um dos programas dos canais concorrentes cuja duração deverá coincidir em 50% com a duração do programa de referência. Com base nos índices parciais construíram-se duas bases de dados para os períodos horários e consideram-se os valores para o universo televisivo. Associou-se aos períodos horários o valor do *rating* global para os programas que cada faixa temporal abrange e construiu-se uma segunda base de dados com os valores de *ratings* parciais. Por vias do *software* Telereport construiu-se uma terceira base de dados com os índices referentes ao canal.

Para as variáveis associadas ao canal ou períodos horários será, também, necessário a construção de uma segunda base de dados simbólica, na qual as variáveis em estudo se referem a períodos horários de 30 minutos de duração, em vez de a valores programáticos. Considera-se a totalidade da emissão diária, mas diferenciando o fim-de-semana dos dias laborais. As variáveis simbólicas são da tipologia intervalo $[\alpha, \beta]$ na qual α representa o limite inferior observado para a variável *Rat_Período%* ou *Share_Período%* durante o correspondente intervalo de tempo e β será o limite superior observado. Os conceitos correspondem a intervalos de tempo (Δt), como, por exemplo, 19H a 19.30H (S) ou (FS). Os períodos horários semanais representam-se pela sigla (S) e os valores de fim-de-semana por (FS). No final agregaram-se a esta base de dados os valores de *ratings* por períodos horários dos diversos segmentos sociodemográficos.

F.1.2.4. Metodologias

As metodologias a utilizar correspondem às que foram utilizadas para a caracterização das taxas médias de audiências para o total da TV.

F.1.2.5. Resultados

I. Caracterização da evolução da variável *Rat_Prog%*

Por análise das séries temporais de valores de *Rat_Prog%* pretendemos verificar se existe um padrão de evolução associado à variável em estudo, ainda que provavelmente dependente de factores temporais ou do canal considerado, uma vez que estes diferem quanto à composição sociodemográfica da sua audiência, posicionamento no mercado televisivo, objectivos estratégicos e quanto à estruturação da grelha de programação e tipologias programáticas transmitidas.

Recorre-se essencialmente a análises gráficas e ao *clustering* de dados simbólicos, onde os objectos de segunda ordem correspondem às diversas transmissões de determinados programas, diferenciados em função do período semanal ou de fim-de-semanam, por exemplo, *espaço_informativo_noite_SIC_semana*.

Atendendo aos gráficos que ilustram a evolução observada nos valores dos *ratings* programáticos diários globais e considerando a totalidade da duração de um programa, verificamos que existe uma certa semelhança na evolução diária desta variável para todos os canais considerados. No entanto, isto não significa, obrigatoriamente, a possível existência de alguns hábitos de exposição televisiva, em grande parte condicionados pelos factores de disponibilidade do painel de audimetria.

A avaliação da possível existência de hábitos de consumo televisivo e do grau de regularidade da exposição será realizado com a análise dos indicadores acumulados de audimetria nos seguintes capítulos. Porém, a evolução das séries temporais corresponde à que condiciona a viabilidade de previsão destes indicadores, tendo sido demonstrado que os dados para o colectivo corresponde à tipologia de dados mais adequada às análises de previsão na área do planeamento de *media*.

Considerando a globalidade dos dados, os canais privados generalistas apresentam um padrão de evolução de *ratings* não só mais semelhante no decorrer das semanas em análise para cada ano considerado, como também comparando os vários anos. O canal estatal RTP1 é o canal que mais se diferencia a ambos os níveis. No entanto, recorrendo a dados de *Rat_Período%* e por aplicação de algoritmos de *clustering* de ST verificar-se-á se estes resultados se mantêm.

Esta maior irregularidade dos níveis de audiências do canal estatal RTP1, também resulta da frequente introdução na grelha de programas de eventos desportivos; numa maior incidência em programas que não são transmitidos em episódios e por se verificarem maiores alterações nas tipologias programáticas transmitidas ao longo dos anos. No período semanal e sobretudo a partir do período designado de pré-prime-time não existe uma estratégia de fidelização das audiências, transmitindo-se continuamente os mesmos programas ou tipologias ao longo dos anos.

Pelo contrário, os canais privados com o intuito de fidelizar o grande público praticam a fidelização horizontal e a vertical. Segundo *Fernandes (2001)*, as grelhas da RTP1, da SIC e da TVI são todas federativas, ou seja, existe a procura dos públicos e cada faixa horária tenta se adequar ao público potencialmente disponível. Ainda assim, perante o horizonte temporal em análise são encontradas diferenças quanto à variabilidade e evolução dos *ratings* entre os canais privados generalistas e a RTP1.

Por exemplo, a RTP1, desde há longa data, transmite o concurso “O Preço Certo”. Nos espaços adjacentes e após o espaço informativo da noite, este canal também aposta nos programas de humor, em concursos, programas informativos (reportagens e debate), eventos desportivos, filmes, programas musicais, entre outras tipologias. Verificamos que até 2010 mantêm-se as mesmas tipologias de programas nos espaços adjacentes aos espaços informativos, com uma forte incidência em concursos ou em programas informativos. No decorrer das manhãs desde 2000 até 2012 transmite-se o programa “Praça de Alegria” e após o espaço informativo da tarde alguma série ou telenovela e talkshows nacionais.

Os canais privados, por sua vez, nos espaços adjacentes ao espaço informativo da noite e nos períodos de maiores audiências apostam essencialmente em programas de ficção, transmitindo telenovelas, séries, programas de humor e os “reality shows” entre outros, concursos, no decorrer da década 2000-2010.

A RTP1 aposta em maior grau por uma concorrência por complementaridade e os canais privados por mimetismo, atendendo às suas necessidades de financiamento comercial.

Para o período em análise as observações atípicas de valores de *ratings* para o canal estatal RTP1, correspondem essencialmente a alguns jogos de futebol da selecção ou dos principais clubes nacionais. Verificamos que os restantes eventos desportivos, sejam estes de futebol ou de outras modalidades desportivas não se repercutem de forma tão significativa na evolução dos *ratings* do próprio canal e dos canais concorrentes para os anos em análise.

Identificando, as observações atípicas da variável *Rat_Prog%*, os programas que apresentam uma maior repercussão são, sem dúvida, essencialmente alguns jogos de futebol. No entanto, estes diferenciam-se quanto ao grau da sua influência, o que se deve não só aos clubes envolvidos, como, também, aos horários em que os jogos são transmitidos e ao canal. No entanto, a maior parte dos jogos, ou pelo menos, os mais relevantes foram transmitidos nos períodos de maiores audiências entre as 19H e as 21H, ou seja, nos períodos que correspondem ao espaço informativo da noite e às faixas horárias adjacentes.

As observações atípicas nos canais privados, também dizem essencialmente respeito a alguns eventos desportivos de maior relevância e a determinadas transmissões relacionadas com acontecimentos sociais e informativos que apresentam um maior destaque. Verificamos que apenas alguns programas incutem uma forte alteração na evolução dos *ratings*.

A introdução de alguns eventos desportivos na grelha de programas poderá alterar significativamente a evolução diária desta variável, assim como os *ratings* dos canais concorrentes. A repercussão observada será função do canal de transmissão, faixas horárias, equipas envolvidas ou da relevância do jogo e do grau de incerteza associado ao seu resultado final, entre outros factores.

No período em análise a maior parte dos eventos desportivos foram transmitidos no canal estatal RTP1. Com base nas análises gráficas verificamos que os canais privados podem ser penalizados, verificando-se consideráveis alterações na evolução dos seus *ratings* quando este estatal transmite programas que podem apresentar uma grande audiência. Por análise das séries temporais multivariadas de valores de *ratings* e de *shares* de audiências dos vários canais verificamos que nestas situações surgem vários cenários possíveis: repercussão apenas ao nível dos *shares* (Cenário A) ou repercussão também ao nível dos *ratings* (Cenário B) e neste último contexto poderá abranger as faixas horárias adjacentes ou não.

As previsões nestes contextos são mais difíceis de realizar, atendendo aos factores mencionados no âmbito dos modelos de previsão das audiências televisivas (*Apêndice D*) e que são novamente abordados no *Capítulo V. A priori* a identificação de cenários semelhantes poderia contribuir para as previsões de *ratings* de alguns eventos desportivos. Contudo, estas abordagens são pouco utilizadas na prática, uma vez que é difícil encontrar no histórico situações muito semelhantes à actual, sobretudo se considerarmos muitas características da

grelha de programação, entre outras, e devido à necessidade de recorrer em maior grau ao passado recente. Se considerarmos um intervalo de tempo mais abrangente os resultados podem perder fiabilidade uma vez que a estrutura dos meios de comunicação e o desempenho dos canais se vai alterando no decorrer do tempo, assim como os padrões de exposição ao meio televisivo que será função das vivências e do quotidiano dos telespectadores.

Cenário A

Por vezes, os *ratings* e os *shares* do canal estatal sobem significativamente, penalizando os canais privados ao nível dos *shares*, mas mantendo-se a mesma ordem de grandeza nos seus *ratings*. Nestas situações o incremento verificado nos *ratings* e *shares* deve-se em princípio ao público que geralmente não vê TV nesses períodos e, não, ao público que poderá ser retirado aos canais concorrentes.

Cenário B

Por oposição, surgem situações em que os *ratings* dos canais privados ficam reduzidos a metade dos valores normalmente observados para idênticos períodos horários, por exemplo, quando se transmitem os jogos Sporting-Benfica ou da Selecção. A extensão desta influência e a repercussão nos programas adjacentes é também variável de programa para programa e dependente das faixas horárias e da contra-programação.

Contudo, por análise da variável *Tvr_Periodo%* verificamos que a sua evolução média é muito semelhante, ano após ano, inclusive considerando períodos que se caracterizam por significativas alterações nos valores de *shares* de audiências dos canais generalistas. Logo, as previsões nos contextos em que se transmitem programas mais populares apresentam uma fiabilidade que é questionável, uma vez que se encontram fortemente dependentes dos factores de disponibilidade e da vontade de ver TV de quem não apresenta uma exposição mais regular ao *media* e das suas eventuais preferências e interesses que não são previsíveis ou dos restantes elementos do agregado familiar ou do grupo em que este se encontra inserido nesse momento. A exposição poderá ser intencional ou não e decorrer da influência do grupo que rodeia cada telespectador e de outras variáveis estruturais e individuais. No capítulo VI consta uma proposta metodológica que pode contribuir para o planeamento de *media*, nos contextos de uma maior irregularidade da exposição do painel. Foi anteriormente mencionado um trabalho de análise de dados de audimetria realizado em 2006 com dados gregos (Assinkopoulo, Goodwin, Nikolopoulos & Patelis, 2006). Os autores, recorrendo ao *Data Mining* tentaram prever a repercussão destes eventos nos *shares* de audiências. O passado poderá ser estruturalmente diferente do momento actual devido às contínuas alterações que se observam no *media* e ao nível da constituição das grelhas. Se nos concentrarmos no passado recente, onde será mais fácil encontrarmos situações mais semelhantes, também poderá não existir um número significativo de casos para se ter êxito nas previsões a realizar. Contudo, esses autores conseguem uma razoável capacidade explicativa da evolução dos dados.

Com base nas análises realizadas na presente tese, verificámos a existência de uma grande diversidade de cenários possíveis. Apenas probabilidades podem ser associadas à direcção mais provável de evolução dos dados.

Aparentemente apenas programas de maior popularidade do canal estatal RTP1 conseguem retirar audiência aos canais concorrentes, influenciando a respectiva evolução dos *ratings*. A repercussão ao nível das galas e de outros eventos nos *ratings* é muito menor, como por exemplo, as galas dos Concursos da Canção da Eurovisão.

Em relação à transmissão de programas que apresentam uma maior popularidade por parte dos canais privados, que não sejam determinados eventos desportivos, a sua repercussão nos canais concorrentes também não é muito significativa para o período em análise. Na presente tese, dar-se-á prioridade à análise das grelhas de programação que não apresentam observações atípicas, uma vez que são as que talvez possam viabilizar melhores resultados do ponto de vista da previsão.

Nos dias em que não se transmitem eventos desportivos, a RTP1 apresenta uma evolução diária mais semelhante com um primeiro pico de audiências que corresponde ao espaço informativo da tarde e um segundo significativamente superior que corresponde ao espaço informativo da noite, ao qual se segue na maioria dos dias um concurso com índices ligeiramente inferiores. Após o espaço informativo da hora do almoço existe uma acentuada redução nos *ratings* invertendo-se esta tendência a partir do concurso “O Preço Certo”.

Após o concurso da noite verifica-se uma vez mais uma acentuada redução nos índices de audiências. Segundo *Fernandes (2001)*, o público da RTP1 é rural e mais idoso. Nos capítulos subsequentes, em especial no *Apêndice F.5* avalia-se o perfil de audiências de alguns canais e tipologias de programas. Sendo verdade a afirmação de Fernandes, apesar da maior disponibilidade deste público, o canal não consegue prolongar durante um grande horizonte temporal o período em que se maximizam os seus níveis de audiências. Portanto, para este canal os espaços informativos funcionam como pilares de audiências, podendo favorecer unicamente os espaços directamente adjacentes. Os espaços informativos, também correspondem aos programas mais seleccionados pelo público deste canal. As galas de programas musicais transmitidos após o espaço informativo da noite apresentam *ratings* da mesma ordem de grandeza, face aos restantes anos para idênticos períodos horários.

A grelha semanal da SIC nos períodos de maiores níveis de audiência apresenta uma maior estabilidade no decorrer dos vários anos em análise. Desde o final da tarde transmite frequentemente telenovelas brasileiras ou nacionais, assim como nos espaços adjacentes ao espaço informativo da noite. Nos espaços adjacentes ao espaço informativo da tarde aposta em programas “*talk show*” e em telenovelas que, por vezes, são retransmitidas. A TVI para os mesmos anos incide a sua programação nas telenovelas de ficção nacionais e em *talk shows* ou *reality shows*. Os canais privados mantêm ao longo dos anos determinadas tipologias de programas nas mesmas faixas horárias.

Assim, incentivam o público a desenvolver comportamentos de lealdade programática, tipológica e ao próprio canal e que se avaliará nas análises de duplicação das audiências a realizar. Apesar da lealdade desenvolvida ao longo dos anos a determinados programas e canais, o conteúdo dos canais concorrentes poderá pontualmente influenciar a selecção programática, alterando certas tendências nos comportamentos dos telespectadores.

Segundo *Fernandes (2001)*, apenas na direcção da SIC se reconhece que no decorrer do tempo os telespectadores desenvolvem lealdade ao canal. As restantes estações televisivas afirmam que os telespectadores seguem os seus interesses e esquecem-se em que canal cada programa é transmitido.

Até 2004 a TVI apresenta uma tendência ascendente nos seus *shares* de audiências, tentando consolidar a sua posição e a partir deste ano apresenta um valor semelhante ao do canal SIC. Frequentemente, a telenovela do horário nobre do canal SIC atinge um máximo de audiências que supera os valores do espaço informativo da noite e que pode favorecer o programa que lhe antecede. Neste contexto são utilizadas as estratégias de *Hammocking*, colocando-se um programa com menores índices de audiências entre dois programas com melhores desempenhos. Ao Domingo, o pico de audiências corresponde ao programa após o espaço informativo da noite. A partir das 17H verifica-se o início de uma tendência ascendente nos valores dos *ratings*, invertendo-se esta tendência após a primeira ou segunda telenovela da noite em função dos anos.

O período ascendente na SIC tem início numa faixa horária muito anterior à da RTP1 e o diferencial de valores de *ratings* entre os dois espaços informativos, também é muito superior ao do canal RTP1. A ordem de grandeza dos níveis de audiência associados ao espaço informativo e às telenovelas é semelhante. Em alguns anos, por exemplo, em 2002 este canal consegue prolongar o seu horário nobre até às 24H, uma vez que transmite um concurso antes da telenovela da noite, mantendo-se a mesma ordem de grandeza das audiências desde as 20H até à conclusão ou no decorrer da mencionada telenovela. No decorrer dos anos através de diferentes estratégias de programação consegue um maior ou menor horizonte temporal de maximização das audiências. Segundo *Fernandes (2001)*, a SIC apresenta um público em que incidem os mais jovens e as donas de casa, mas no horário nobre todos os segmentos estão presentes. O canal TVI é associado aos jovens entre os 15 e os 40 anos e das classes mais altas (AB e C1).

Para o canal TVI, as observações atípicas também correspondem essencialmente a alguns eventos desportivos. Por vezes, no período de dois meses considerado em cada ano, identificamos apenas um número muito reduzido de valores de *ratings* que se diferenciem significativamente. As telenovelas do horário nobre tendem a superar as audiências do espaço informativo da noite. A tendência ascendente tem início a partir das 18H e a descendente a partir das 22H. Em alguns anos, por exemplo, em 2003, os maiores níveis de audiências mantêm-se até à conclusão da segunda telenovela por volta das 24H.

O diferencial de *ratings* entre os dois espaços informativos é muito significativo e corresponde ao maior valor observado comparativamente aos canais SIC e RTP1 se considerarmos o período em que a TVI apresenta uma gradual tendência crescente no valor do seu *share* de audiência.

Por último, para o Canal 2 os *ratings* são irrisórios, subindo ligeiramente para o espaço informativo da noite e, por vezes, para o seu programa sucessor. Os índices de audiências são inferiores aos 2.5% o que torna este canal pouco atractivo do ponto de vista comercial. Este canal destina-se a um público minoritário, com uma grelha de programação compósita na qual o responsável da programação decide o que mostrar ao público e não existe uma construção da grelha que vise a maximização das audiências televisivas.

II. Análise da Variabilidade da Variável *Rat_Prog%*

A construção de uma base de dados simbólica permite-nos calcular um intervalo de variabilidade de valores de *Rat_Prog%*, por exemplo, agregando todos os espaços informativos de determinado canal, independentemente do seu horário de início. Pretende-se tentar avaliar se existe alguma estabilidade nos valores dos *ratings*, independentemente das suas alterações horárias ou ao nível das características da contra-programação e da programação adjacente. Por aplicação de algoritmos de *clustering* simbólicos identificamos quais são os programas que apresentam uma maior variabilidade de valores.

Uma vez que os padrões de consumo televisivo se diferenciam entre o fim-de-semana e os dias laborais, os programas que são transmitidos em ambos os períodos são identificados pelo período a que correspondem. Se a variabilidade for significativa, em princípio poderá significar que prevalece a dimensão temporal e que, consequentemente as previsões devem ser realizadas em função das faixas horárias em que os programas são transmitidos, em vez de a nível programático. Pretende-se, essencialmente, avaliar o grau de independência do desempenho programático face às variáveis temporais e de contra-programação.

Analisando a variabilidade observada nos valores dos *ratings* para programas em concreto verificamos, por exemplo, que o espaço informativo da noite da TVI (S) apresenta horários de início entre as 19.30H e as 21.30H e ao fim-de-semana entre as 20H e as 21.30H, portanto não transcendendo o período designado de prime-time ou de pré-prime-time. Apresenta margens de 10,3% de variabilidade ao fim-de-semana e de 7,1% para os dias laborais. A margem de 10,3% resulta deste programa ter sido favorecido com a prévia transmissão da Final da Taça de Portugal. Excluindo apenas este programa, a variabilidade é menor com o valor de 6,7%, o que representa um decréscimo de 3,6%. Em relação ao período semanal, a variabilidade é acentuada com a redução dos *ratings* quando se verifica a transmissão simultânea de alguns jogos de futebol. Se excluirmos estes programas a variabilidade desce para 4,5%, o que representa um decréscimo de 2,6%.

Os programas desportivos adjacentes ou concorrentes contribuem para a variabilidade observada, no entanto excluindo-os os valores de variabilidade ainda permanecem com valores consideráveis, atendendo aos valores dos *ratings* e em especial ao fim-de-semana.

Verificamos que a SIC apresenta uma posição mais consolidada no mercado televisivo no decorrer do total do período em análise. Consequentemente, o seu espaço informativo da noite é menos variável nos seus valores de *ratings*, o que podemos atribuir, quer a uma maior estabilidade horária, quer à “lealdade” do público para com a telenovela da noite e para com o canal o que impulsiona a ver os seus programas antecessores (4% de variabilidade ao fim-de-semana e 6,6% durante a semana). Aparentemente neste canal é superior a variabilidade de *ratings* semanais.

Se excluirmos os programas que apresentam como contra-programação os jogos de futebol do canal estatal, a variabilidade dos valores de *ratings* durante a semana reduz-se para 3,4% (-3,2%) e ao fim-de-semana para 1,3% (-2,7%). Logo, uma maior disponibilidade do painel reduz a variabilidade observada nestes indicadores. A imagem de um canal, também contribui para uma maior estabilidade dos seus valores em períodos de lazer.

Os horários de início dos espaços informativos da noite para o canal estatal são consideravelmente variáveis, com a frequente transmissão de jogos de futebol e de galas em directo nos períodos adjacentes o que influencia os seus intervalos de valores de *ratings* (12% ao fim-de-semana e 17% durante a semana) e, por vezes, como foi demonstrado, também, dos canais concorrentes.

Uma vez mais ao fim-de-semana, também apresenta uma maior estabilidade de valores, assim como se verifica no canal SIC. Se excluirmos os eventos desportivos os valores de variabilidade reduzem-se para 3,9% durante a semana e para 2,2% ao fim-de-semana. A repercussão do futebol e, por vezes, do festival da canção é muito significativa. Neste último caso, essencialmente, como consequência da alteração horária que se verifica neste programa e não como uma função das alterações no número de telespectadores.

Podemos observar que um pico de audiências para os espaços informativos da noite do canal estatal pode corresponder a um forte declínio nos programas dos canais privados. Os *ratings* do espaço informativo da noite da RTP1 são frequentemente superiores e os valores correspondentes para os canais privados alternam-se no decorrer do período em análise.

Excluindo determinados programas, a RTP1 e a SIC correspondem aos canais que apresentam uma menor variabilidade de valores de *ratings*, uma vez que as suas posições encontram-se mais consolidadas no mercado televisivo para o horizonte temporal em análise.

Os valores de *ratings* dos programas do Canal 2 são irrisórios e pouco variáveis.

Em suma, os *ratings* dos espaços informativos da noite encontram-se essencialmente influenciados por alterações horárias, canal de transmissão, período semanal ou de fim-de-semana e pela programação adjacente e concorrente.

Os espaços informativos da tarde apresentam índices mais reduzidos, mas muito mais estáveis, atendendo à amostra global. À hora do almoço em princípio existe uma audiência mais regularmente disponível, mas obviamente de menor dimensão. Com base nas análises de duplicação de audiências confirmar-se-á o grau de regularidade existente ao nível da exposição programática. Os valores de variabilidade de *ratings* dos espaços informativos da tarde da RTP1 são de 1,8% para os programas semanais e de 5,7% ao fim-de-semana, sendo estes valores de 1,4% (programa semanal) e 2,8% (programa de FS) para a TVI ou de 2,7% (S) e 6,3% (FS) para a SIC.

Os valores semanais de variabilidade dos espaços informativos da tarde são inferiores aos valores de fim-de-semana, sendo o diferencial superior para o canal RTP1 e para a SIC. Por exemplo, ao fim-de-semana a variabilidade é de 0,7% para a RTP1. Se excluirmos o espaço informativo da SIC que teve como contra-programação uma transmissão de Fórmula 1 do canal estatal RTP1 a variabilidade reduz-se para 3,2% (6,3% era o valor inicial). Neste caso o horário do espaço informativo da tarde (RTP1) foi antecipado e parte do seu público viu em substituição o programa da SIC, maximizando ligeiramente os valores dos *ratings* para este canal e penalizando a RTP1. Os valores semanais de *ratings* apresentam variações no máximo de 2,7%, mas é necessário ter em atenção que os respectivos valores de *ratings* também apresentam valores reduzidos, e, logo a nível relativo a variabilidade pode ser considerada como sendo considerável.

Em suma, para os espaços informativos da tarde, a influência da contra-programação e das alterações horárias é notória em especial ao fim-de-semana.

Se considerarmos apenas a grelha de programas normalmente transmitida existe uma certa estabilidade nos valores dos *ratings*. A variabilidade observada não é muito significativa atendendo à dimensão do painel. A fidelidade que o público dirige aos espaços informativos é uma variável a ter em atenção, uma vez que influencia os programas adjacentes. Os programas semanais apresentam uma maior estabilidade de valores.

Considerando os programas do horário nobre e da noite que maximizam os *ratings* para os canais correspondentes, verificamos que a sua popularidade reduz a variabilidade observada nos seus *ratings*.

A lealdade que o público aparenta dirigir aos programas mais populares de cada canal favorece uma maior regularidade nestes indicadores. Maiores valores dos *ratings* programáticos acarreta uma menor variabilidade, no entanto apenas analisando os índices de lealdade programática (*Apêndice F.3*) podemos constatar se no decorrer do tempo, existe alguma estabilidade na composição da sua audiência e ao nível da regularidade da exposição.

A principal telenovela da noite varia o seu horário de início entre as 21.30H e as 24H e a segunda telenovela tem início entre as 22H e as 24.30H.

Logo, é inevitável a existência de uma maior variação nos valores de *ratings* destes programas dada a evolução polinomial que se verifica nestes períodos para a variável *Tvr_Período%*. Apesar da popularidade de alguns programas, as alterações observadas nos seus horários não retêm a audiência o suficiente para poder existir uma maior estabilidade de valores. Independente da popularidade de um canal ou de um programa a influência do factor temporal é determinante. Para todos os programas a variabilidade dos *ratings* é menor ao fim-de-semana.

O horário da principal telenovela da noite é ligeiramente menos variável ao fim-de-semana, verificando-se entre as 22H e as 23.30H e para a telenovela seguinte situa-se entre as 22H e as 24H. Uma maior disponibilidade do público é um dos factores que justifica a existência de uma maior regularidade nos valores de *ratings* ao fim-de-semana (2,9%). Em dias de lazer mais facilmente o público poderá esperar pelos programas que lhe despertam um maior interesse, atendendo aos conteúdos disponíveis.

O concurso da noite da RTP1 apresenta uma menor variabilidade de valores (3,6%), o que poderá ser atribuído à composição sociodemográfica da audiência do canal estatal e à sua maior estabilidade horária.

O concurso “O Preço Certo” apresenta 3,9% de variabilidade durante os dias laborais e de 6,3% ao fim-de-semana. No período de pré-primetime a variabilidade dos *ratings* é superior ao fim-de-semana.

As telenovelas da SIC, também não apresentam grandes intervalos de valores de *ratings*, sendo estes minimizados ao fim-de-semana.

Para a primeira telenovela da noite da SIC, os valores são de 4,3% (programa semanal) e 2,2% (programa de FS); para a segunda telenovela de 3,8% (S) e 1,2% (FS) e para a terceira de 2,3% (S). Quanto mais avançamos no decorrer da noite menor será a variabilidade e a dimensão dos valores de *ratings*. O programa de humor da noite da SIC é muito variável nos seus horários apresentando intervalos de variabilidade de 8,8% (S) e de 12,9% (FS).

A variabilidade ao nível dos *shares* é mais significativa, uma vez que esta variável é fortemente influenciada por variações nas taxas médias para o total da TV. Uma maior estabilidade nos valores dos *ratings* poderá não ser acompanhada ao nível dos *shares*.

Com base nas análises gráficas foi possível verificar a existência de um aparente padrão diário de evolução de *ratings* e constatar que os programas apresentam, por vezes, uma variabilidade considerável ao longo das semanas em análise. Isto atribui-se essencialmente às alterações verificadas nos seus horários o que condiciona o valor das taxas médias de audiências para o total da TV e, por vezes, à contra programação a eles associada. As alterações mais significativas verificam-se quando a contra- programação inclui os eventos desportivos. A introdução de um programa com um maior desempenho por parte de um dos canais, inevitavelmente, também, altera a distribuição dos *shares*, repercutindo, por vezes, também, nos seus programas adjacentes.

Quando o canal estatal transmite um jogo de futebol no decorrer da tarde ou no horário nobre; programas de Fórmula I no horário do almoço, ou alguma gala, como, por exemplo, o festival de Eurovisão no horário nobre, os

restantes canais podem ser penalizados, não só, nos mesmos períodos horários, como, também, nos subsequentes devido aos efeitos de herança que a popularidade destes programas pode gerar. Esta influência é mais notória quando os espaços informativos da noite são antecidos por eventos desportivos.

Os canais concorrentes pontualmente, também podem ser beneficiados, especialmente quando existe uma antecipação do horário do espaço informativo da RTP. O horário nobre é um dos períodos que apresenta uma maior irregularidade ao nível das taxas médias para o total da TV, consequentemente neste período os *ratings* também são mais variáveis, especialmente se considerarmos os programas menos populares aos quais não se possa associar comportamentos de lealdade.

Nos períodos que apresentam maiores níveis de audiência também são mais frequentes as alterações na grelha de programas, não só, a nível qualitativo, como, também, quanto aos seus horários de transmissão. As tardes e as manhãs dos dias laborais, em princípio devem apresentar uma maior estabilidade nos seus níveis de *Tvr_Prog%* e consequentemente ao nível dos *ratings*.

Com base nos resultados apresentados verificamos que a afectação de *Tvr_Periodo%* a cada um dos canais pode variar ao longo dos anos. No entanto, basear as análises no desempenho de meras tipologias programáticas é demasiado redutor e dificilmente se conseguirá prever com exactidão como o público reagirá à introdução de determinado programa. Os resultados devem ser reavaliados quando se verificam consideráveis alterações na grelha.

Logo, o objectivo das análises traduz-se em caracterizar a dinâmica das escolhas realizadas, ou seja, na identificação de tendências. Os resultados apresentados resultam de considerar índices programáticos globais.

Considerando os dados parciais deverá existir uma menor variabilidade, condicionada pelos períodos horários a que dizem respeito. Os valores globais podem não resultar de tendências uniformes no decorrer de todo o programa, mas apenas para algumas fracções do mesmo. As audiências programáticas parciais são mais relevantes do ponto de vista comercial. No entanto, não devemos esquecer que o comportamento da audiência no decorrer de um intervalo pode estar influenciada, para além dos factores temporais, também por parte da durabilidade programática global a jusante e não só a parcial. Consequentemente, também é necessário a análise dos *ratings* a um nível mais desagregado.

Com base nas análises realizadas verificamos que existe uma forte diversidade de cenários possíveis e identificámos quais são os factores responsáveis por uma maior variabilidade de *ratings* e em que direcção se manifesta.

Segundo alguns autores, a regularidade da grelha de programas constitui a melhor forma de criar hábitos de visionamento que, na sua perspectiva constitui a chave de sucesso de um canal. Verificamos que ainda assim a variabilidade dos *ratings* pode ser muito significativa quando se verificam alterações horárias na sua transmissão. Logo, no seguinte item considera-se os valores de *ratings* associados a faixas horárias.

III. Análises Exploratórias da Evolução da Variável *Rat_Período%* e *Shr_Período%*

Uma vez que os valores globais de *ratings* associados a programas em concreto reflectem também as consequências da variabilidade observada nos seus horários, interessa verificar se existe uma maior estabilidade nestes valores se considerarmos períodos horários, em vez de programas. Verificámos que a alteração de um programa para diferentes faixas horárias condiciona significativamente os valores dos seus índices de audiências.

Dividiu-se a emissão diária em intervalos de meia hora e associou-se a cada intervalo a média ponderada dos *ratings* programáticos parciais associados a cada um dos períodos considerados. Esta fórmula de cálculo é uma simplificação, uma vez que os *ratings* reflectem as taxas médias de audiência atendendo à duração global de um programa e não para os intervalos de meia hora. Para a variável *Tvr_Prog%* verificámos que os erros em que se incorre são reduzidos, uma vez que existe uma grande semelhança de valores para períodos adjacentes. Alternativamente, podemos calcular o valor dos *ratings* associado ao canal para o mesmo intervalo de tempo. Para a variável *share* os valores referentes às faixas horárias foram calculados da mesma forma.

Do ponto de vista metodológico recorreu-se a análises gráficas e às metodologias de *clustering* de séries temporais enumeradas anteriormente, ou seja, o cálculo da distância euclidiana e de coeficientes de correlação com a posterior aplicação do algoritmo K Média.

Por observação dos gráficos que traduzem os valores médios, máximos e mínimos associados a cada um dos canais para os valores de *ratings* e de *shares* de cada intervalo de meia hora desde as 9H até às 24.30H foi possível retirar as conclusões que se descrevem nos seguintes parágrafos.

Para o canal estatal RTP1 os valores máximos de *ratings* aumentam significativamente entre as 19H e as 23H, sendo estes também os períodos onde se observa uma maior amplitude de valores (<20%) e de valores médios. A restante emissão diária apresenta uma evolução média muito próxima dos valores mínimos observados e de reduzida dimensão e variabilidade (<5%).

A variável *Shr_Período%* apresenta em grande parte do dia uma significativa amplitude de valores. Os valores mínimos e máximos de *shares* da RTP1 oscilam significativamente no decorrer do dia, sendo superiores a partir das 19H. Esta variável apresenta uma grande amplitude de valores e que ultrapassa a observada para os valores de *ratings*, sendo aparentemente mais difícil de prever. Os valores de *shares* médios apresentam uma maior regularidade no período matinal e no decorrer da tarde, decrescendo linearmente a partir do horário nobre.

Para o canal TVI os limites de variabilidade de *shares*, também são muito irregulares, podendo ultrapassar os 30% de diferencial de valores para idênticos períodos horários. Para o período em análise os valores são superiores a nível médio a partir das 19H, observando-se uma evolução crescente do tipo linear.

No decorrer do período matinal e da tarde existe alguma regularidade na evolução média deste indicador.

Os *ratings* do canal TVI, também apresentam uma amplitude de valores muito inferior aos valores de *shares*, maximizando-se a partir do horário nobre. Os valores extremos dos *ratings* distanciam-se de forma máxima entre as 20H e as 22H. Independentemente das alterações observadas na programação, os valores médios de *ratings* apresentam uma certa regularidade até às 20H. Será a partir do espaço informativo da noite que se justifica em maior grau recorrer a modelos multivariados das audiências televisivas.

O canal SIC apresenta maiores intervalos de variabilidade de valores de *ratings*. Os *ratings* médios também apresentam uma menor consistência no decorrer do dia, maximizando-se para um maior horizonte temporal que tem início às 16.30H. Os seus valores de *shares* apresentam a mesma irregularidade, mas nos limites máximos de forma menos acentuada em relação aos canais concorrentes.

No entanto, a SIC corresponde ao canal que apresenta um maior horizonte temporal com maiores valores máximos de *shares* e de valores médios deste indicador. O canal estatal apenas é beneficiado quando transmite eventos desportivos. A sua grelha de programação, no geral considerada mais interessante por parte do painel, permite uma maior estabilidade ao nível de *shares*, apesar das alterações que se possam verificar em *Tvr_Período%* e ao nível da estruturação da grelha de programas.

Verificamos com base nestes resultados que os canais divergem quanto às faixas horárias em que apresentam diferenciadas evoluções de *ratings* e que este indicador acaba por ser menos variável do que os valores dos *ratings* programáticos. Por sua vez, os valores de *shares* apresentam uma variabilidade ainda superior não devendo *a priori* constituir uma variável explicativa ou de previsão dos valores de *ratings* em análises multivariadas a realizar.

IV. Análise de Valores de *Ratings* Sociodemográficos

Os *ratings* programáticos para o universo televisivo encontram-se obviamente dependentes da sua evolução diária para os vários segmentos sociodemográficos representados nas bases de dados de audimetria. Os *ratings* globais resultam da exposição individual e sociodemográfica, a qual será variável no decorrer de um dia e da semana.

Consequentemente, interessa verificar quais são os segmentos que se aproximam em maior grau entre si e face ao universo televisivo, considerando a totalidade da emissão diária e determinadas faixas horárias. Também devemos verificar se a variabilidade observada nos dados para a totalidade do painel pode ser atribuída a um reduzido número de segmentos e se estes serão os mesmos ou não, ao longo do dia e da semana.

Com o intuito de desenvolver modelos explicativos e de previsão de audiências, os segmentos que apresentam uma maior variabilidade ao nível do consumo televisivo

devem ser objecto de análises mais específicas que promovam a avaliação da repercussão de variáveis estruturais e individuais na sua exposição televisiva. O peso de cada um destes segmentos deve ser considerado na definição da composição sociodemográfica das faixas horárias associadas a determinada programação e da sua contra-programação. Pretende-se, também, identificar quais são os períodos horários e os canais que maximizam os *ratings* para cada um dos segmentos representados nas bases de dados disponibilizadas.

A variável em estudo corresponde aos *ratings* sociodemográficos para a totalidade da emissão diária. Os *ratings* traduzem as taxas médias de audiências associadas a programas em concreto, considerando durações globais e parciais. Nas análises simbólicas a realizar, os objectos simbólicos booleanos representam para cada segmento os intervalos de variabilidade de *ratings* para cada intervalo de meia hora desde as 7H até às 24H. Por exemplo, ao intervalo horário [7.30H-8.00H] corresponde várias variáveis da tipologia [a,b] que representam os limites inferior e superior dos *ratings* para determinado segmento.

Com o intuito de avaliar, o grau de semelhança existente entre as séries temporais referentes aos diversos segmentos recorre-se a medidas de dissemelhança e *clustering* clássicas e simbólicas, como, por exemplo, o K Média. Aplica-se também análises Anova para avaliar a repercussão de determinadas variáveis nominais nos índices de audiências.

IVa. Avaliação do grau de semelhança existente entre os *ratings* sociodemográficos com base na análise de séries temporais

Numa fase inicial calcularam-se os coeficientes de correlação com posterior aplicação do algoritmo K Média ou calcularam-se as distâncias euclidianas entre as séries temporais para a totalidade do horizonte temporal considerado. Diferenciaram-se os dias laborais dos dias referentes ao fim-de-semana.

A Distância Eucladiana representa-se por

$$D(Q,C) = \sqrt{\sum_{i=1}^N (q_i - c_i)^2}$$

O Coeficiente de Correlação representa-se por

$$r_{ij} = \frac{\sum_{t=1}^L p_i(t) P_j(t)}{\sqrt{\sum_{t=1}^L p_i(t)^2 \sum_{t=1}^L P_j(t)^2}}, 1 \leq j \leq N, 1 \leq i \leq M$$

Segundo as análises realizadas às séries temporais de valores de *ratings* do canal SIC, verificamos que a classe social C2, o sexo feminino, todas as regiões com excepção do Porto e do Interior, os lares s/ TV Cabo (correspondem aos segmentos mais representados nas bases de dados) e as faixas etárias não extremas (todos os telespectadores excluindo 4-24 anos e a partir dos 55 anos) correspondem aos segmentos que apresentam uma maior semelhança face ao universo televisivo, isto é, do comportamento global da audiência. Consequentemente, para a previsão dos *ratings* deve-se ter em atenção os segmentos que se diferenciam em maior grau, ou seja, os homens,

a classe social AB e D, as faixas etárias extremas, os lares com TV Cabo e as regiões do Porto e do Interior. Em análises de Regressão deve-se ter em atenção as variáveis histograma que representem a distribuição de frequências destes segmentos para os períodos horários em análise. Os segmentos que mais se diferenciam da evolução global diferem de canal para canal.

Consequentemente, em função das características da contra-programação a composição sociodemográfica dos correspondentes períodos horários de cada um dos canais influencia a sua repercussão na evolução dos *ratings*. A atitude destes segmentos face aos espaços publicitários também poderá influenciar as taxas médias de audiências. Verificamos que ao fim-de-semana em geral é maior a aproximação existente entre a evolução dos dados para os vários segmentos e a que caracteriza o universo televisivo, como consequência de uma maior disponibilidade do público. Em dias laborais mais facilmente as características sociodemográficas influenciam os padrões de disponibilidade e consequentemente o processo de selecção programática, podendo alguns segmentos apresentar evoluções mais atípicas.

Com base nas análises realizadas às ST de valores de *ratings* para o canal TVI as faixas etárias que aparentemente mais se distinguem do universo correspondem no decorrer da semana e com maiores valores às idades superiores aos 55 anos e com menores valores os telespectadores que apresentam entre 25 e 34 anos.

Ao fim-de-semana obviamente os mais jovens (<14 anos) apresentam valores superiores à média sobretudo no período da manhã destinado aos programas juvenis. A classe social que mais se diferencia corresponde à classe social D com maiores valores face à média e a classe social AB com valores inferiores, mas de forma mais acentuada no decorrer dos dias laborais. A evolução para o Universo televisivo ocupa um lugar intermédio entre ambos os sexos ao longo de toda a semana. As regiões do Sul e do Interior para este canal diferenciam-se pela positiva no decorrer da semana.

Verificamos que os factores de disponibilidade/vontade de ver TV são os que parecem condicionar em maior grau a dimensão das audiências e o grau de semelhança da evolução dos *ratings* para os vários segmentos. Os segmentos que mais se diferenciam da evolução global correspondem aos que podem apresentar maiores ou menores índices de exposição por ausência ou não de uma actividade profissional.

Para o canal estatal RTP1 verificamos que ao fim-de-semana os diversos segmentos apresentam uma tendência muito próxima à do universo televisivo, diferenciando-se ligeiramente apenas os telespectadores com idades superiores aos 55 anos e com valores superiores à média. Durante a semana diferencia-se ligeiramente a classe D, as regiões do Interior e do Sul e as faixas etárias superiores aos 55 anos com valores superiores à média. Os mais jovens até aos 24 anos ou as faixas etárias no intervalo 35-44 apresentam valores inferiores à média. A faixa etária corresponde à variável que parece condicionar em maior grau a selecção programática para este canal.

Se em vez de considerar a totalidade da amostra, restringirmos as análises apenas aos programas de maior popularidade deste canal verificamos que atendendo aos jogos de futebol que apresentam uma maior repercussão nas audiências (todos os Jogos com um horário de início entre as 19H e as 20H), os intervalos de variabilidade de *ratings* atendendo às variáveis sociodemográficas são os seguintes: o sexo Masculino apresenta valores superiores ao Feminino, com um diferencial de valores entre os 5-10%; as Classes sociais D e C2 apresentam valores superiores às restantes (+5-10%), as faixas etárias representadas nas bases de dados ordenam-se por ordem crescente de *ratings* (10-15%) e entre as diversas regiões geográficas existe entre 5 e 15% de variabilidade de valores. A maior variabilidade de *ratings* está associada às faixas etárias e às regiões. O Porto diferencia-se para os jogos que envolvem o Futebol Clube do Porto.

IVb. Avaliação do grau de semelhança entre *ratings* sociodemográficos (Análises Simbólicas)

Para confirmar as tendências apontadas anteriormente quanto ao grau de similitude existente na evolução dos *ratings* sociodemográficos devemos recorrer a metodologias simbólicas que permitam comparar a evolução de séries temporais atendendo ao grau de variabilidade observado na variável em estudo. Recorremos ao algoritmo de *clustering* Sclust e a medidas de dissemelhança simbólicas. Diferenciam-se as análises em função dos canais e dos períodos horários e consoante nos referimos ao período semanal ou de fim-de-semana.

Iniciam-se as análises recorrendo às medidas de (di)semelhança simbólicas aplicadas a objectos simbólicos booleanos que representam para cada segmento os intervalos de variabilidade de *ratings* para cada intervalo de meia hora desde as 7H até às 24H apenas ao fim-de-semana e para o canal SIC. Posteriormente, comparam-se os resultados por aplicação do algoritmo de *clustering* simbólico Sclust.

□ Resultados para o canal SIC

As medidas de semelhança utilizadas correspondem a U1, U3, U4, SO4 permitindo a construção de mapas que ilustram o grau de proximidade existente entre os diversos segmentos (*Gráficos F.1.2.5.IV.1 e F.1.2.5.IV.2*). Nos contextos em que os resultados foram mais contraditórios e como alternativa também se recorreu às seguintes medidas:

Medida de dissemelhança de Gowda-Diday representada por

$$D(A, B) = \sum_{j=1}^p [D_{1j}(A_j, B_j) + D_{2j}(A_j, B_j) + D_{3j}(A_j, B_j)]$$

Medida de dissemelhança de Ichino-Yaguchi representada por

$$\phi_j(A, B) = |A_j \oplus B_j| - |A_j \otimes B_j| + \gamma(2|A_j \oplus B_j| - |A_j| - |B_j|)$$

Onde $|A_j|$ corresponde ao comprimento do intervalo A_j e $0 \leq \gamma \leq 0,5$ a uma constante pré especificada. $A_j \oplus B_j$ corresponde à união de dois conjuntos e $|A_j \otimes B_j|$ à sua intersecção.

Existe alguma divergência quanto ao posicionamento dos segmentos para as várias medidas de semelhança que foram consideradas, mas os resultados coincidem ao nível dos segmentos que mais se diferenciam dos restantes.

✓ Resultados das análises realizadas com os dados de fim-de-semana

Os segmentos que mais se distanciam dos restantes considerando os dados de fim-de-semana correspondem à classe social A/B, a classe social D, as faixas etárias 4/14, 15-24, 55/64, +64, o Grande Porto e os lares c/ TV Cabo. Portanto correspondem aos segmentos que minimizam (classe social A/B, lares c/ TV Cabo, 15-24) ou maximizam os *ratings* programáticos (Classe social D, 4/14, >55anos, Porto) e incluem as classes sociais e faixas etárias extremas, assim como os lares com um maior número de opções e a região do Porto. O Universo neste contexto está mais próximo dos seguintes segmentos: adultos, não Casa, dona de casa e Classe social C2.

Foi possível estabelecer uma maior similitude entre os seguintes grupos de segmentos de telespectadores:

- Faixas etárias 55-64 e +64. Portanto, a partir dos 55 anos existe uma certa semelhança no valor dos *ratings*;
- Faixas etárias 25-34 e 35-44, sexo feminino, lares sem TV cabo, as regiões do Interior e Litoral Norte, dona de casa e classe social C2. Corresponde a faixas etárias de adultos mais jovens, ao sexo feminino e responsável pelas compras do lar, lares com menor número de canais, a classe social mais baixa e as regiões do Interior ou Litoral Norte;
- Masculino, classe social C1, Lisboa, Litoral Centro e o Porto. Corresponde aos maiores centros populacionais e a uma maior classe social e ao sexo masculino.

✓ **Resultados com a Totalidade dos Dados**

Aplicando o algoritmo SCLUST à base de dados simbólica por forma a identificar *clusters* de segmentos foi possível chegar aos seguintes resultados:

- Os segmentos classe social D, lares s/ TV cabo, Interior e Litoral Norte, +64/ 55-64/ 4-14 são os segmentos que mais se diferenciam dos restantes por apresentarem maiores índices;
- Considerando apenas o período da manhã até às 13H destaca-se a classe social D, o Interior, Litoral Norte, faixa etária 4-14, lares sem TV_Cabo;
- Para o período que decorre entre as 13H e as 18H os *clusters* estão constituídos por: Universo, classe social C2, classe social D, Lisboa, Interior, Litoral Norte, feminino, 55-64, + 64, lares sem TV cabo, donas de casa, não donas de casa, adultos e o segundo cluster pelos restantes segmentos: classe social A/B e C1, faixa etária 4-45, lares sem TV Cabo, Litoral Centro e Sul e sexo masculino;
- Se consideramos os períodos com maiores índices de audiência, isto é, entre as 19H e as 23.30H estão incluídos no mesmo agrupamento os segmentos que apresentam maior exposição televisiva: a classe social D, Lisboa, Porto, Interior, Sul, sexo feminino, faixas etárias 45-55 e +64, donas de casa e lares sem TV cabo.

✓ **Resultados com os Dados Semanais**

Repetindo as análises apenas com os dados semanais os segmentos 4/14, Classe D, 55-+64, A/B, c/ TV cabo diferenciam-se em maior grau dos restantes. Em relação ao fim-de-semana excluímos a faixa etária dos 15-24 e o Porto.

Segundo os mapas de dissemelhanças existe uma maior semelhança entre os seguintes segmentos:

- Os Lares sem TV Cabo, o Interior, o Litoral Norte, feminino e dona_casa;
- Universo, classe social C2, dultos (neste caso não inclui dona de casa e não DC);
- Classe social C1, Lisboa, 35- 44 e 25-34.

Para o período semanal e por aplicação da medida U4 os segmentos que mais se distanciam no global são o sexo masculino, as faixas etárias entre os 4 e os 24 anos e superiores aos 55 anos, as classes AB e D, os lares sem TV cabo e as Donas de casa. Os segmentos que mais se aproximam do Universo são a classe C2 e as idades compreendidas entre os 35 anos e os 45 anos. Restringindo as análises ao prime-time e também por aplicação da medida U4 é também a classe C2 a que se aproxima em maior grau do Universo e os segmentos com evoluções mais distintas são o sexo masculino, as classes AB e D, os adultos, as Donas de Casa, a região do Sul, lares sem TV cabo e as faixas etárias até aos 24 anos.

□ **Resultados para o Canal TVI**

Para o canal TVI verificamos que ao fim-de-semana e por aplicação das medidas de semelhança simbólicas U4, U2, SO1 existe uma maior dispersão dos segmentos relativamente à SIC. Os segmentos que mais se diferenciam dos restantes correspondem a faixas etárias superiores aos 55 anos e menores de 24 anos, a classe social D e a classe social AB, o sexo feminino e donas de casa, regiões do Sul, Interior, Lisboa e Porto. O segmento com uma evolução mais semelhante à do Universo corresponde à classe social C2.

Existe uma maior proximidade entre os segmentos: Interior, feminino, Donas de Casa, Classe D e idades superiores aos 55 anos e entre a classe A/B, masculino e menores de 24 anos.

Restringindo as análises aos períodos que decorrem entre as 19H e as 23.30H e também por aplicação de medidas de dissemelhança os segmentos mais próximos do Universo são a classe C1 e a classe C2. Os segmentos com evoluções mais dispares do conjunto são as faixas etárias até aos 34 anos e superiores aos 55 anos, a classe D, a classe A/B e a Região do Sul.

□ Resultados para o canal RTP1

Para o canal estatal confirmam-se os resultados das análises exploratórias. Verificamos em geral uma maior similitude entre a evolução dos *ratings* para os diversos segmentos.

IVc. Associação de Canais e Períodos Horários a Segmentos Sociodemográficos

Com base nas análises anteriores avaliou-se o grau de proximidade existente entre as evoluções dos *ratings* dos vários segmentos representados nas bases de dados, independentemente da dimensão dos valores. Pretende-se também associar intervalos de *ratings* a segmentos em função dos períodos horários e dos canais, por forma a tentar identificar para cada um deles quais são as associações entre canal e período que permitem maximizar os valores dos *ratings* e diferenciar cada um dos subsegmentos a este nível. Realizaram-se com este objectivo análises Anova.

Segundo uma análise Anova a três factores verificamos que o Canal 2 apresenta menores valores de *ratings* para ambos os sexos independentemente do período horário e sem grandes flutuações. A evolução diurna dos índices de audiências do canal RTP1 também é mais homogénea, superando ou igualando-se a nível médio aos dois canais privados para os períodos horários que decorrem até ao pré-prime-time atendendo ao sexo masculino ou até à tarde para o sexo feminino. Para ambos os sexos existe uma tendência ascendente desde o pré-prime-time que culmina no horário nobre, reduzindo-se ligeiramente no período da noite para o sexo feminino e de forma mais acentuada para o masculino. Neste período a TVI supera a SIC no horário nobre e no decorrer da noite.

O final da noite da RTP1 tem maiores valores de *ratings* do que o período equivalente dos canais privados para ambos os sexos. A diferença entre os valores dos *ratings* médios de pré-prime-time para ambos os sexos é significativa.

Para as classes sociais, segundo a análise Anova realizada, observamos um nítido acréscimo nos índices de audiência com a redução da classe social, em especial desde as 19H até às 24H. Todas as classes apresentam um padrão de variabilidade de *ratings* programáticos muito semelhante, com uma acentuada subida a partir das 19H, atingindo o seu máximo no prime time. Os índices reduzem-se ligeiramente no período da noite e posteriormente decaem acentuadamente. A classe AB apresenta uma maior oscilação diária dos *ratings* médios para o Canal 2. Com a redução da classe social mais se vai acentuando a diferença entre os valores de *ratings* dos canais privados e estatais.

No prime time os canais privados maximizam as respectivas audiências e à hora do almoço apenas para a classe social D sobressai o canal SIC.

Para as faixas horárias e com base nas análises Anova verificamos que a partir dos 55 anos produz-se um acréscimo no valor do pico de audiências associado à hora do almoço, prime-time e no decorrer da noite com uma menor descida de valores dos *ratings* nos períodos intermédios. Também, é notório o acréscimo verificado nas audiências para o canal estatal RTP1. Para todas as faixas etárias o Canal 2 apresenta menores valores de *ratings*.

V. Considerações sobre a análise de valores de *Shares* de audiências

Apesar da variabilidade observada nos *shares*, quando se optar recorrer a este índice para tentar prever os valores dos *ratings* podemos, por exemplo, recorrer às Árvores de Regressão, inclusive simbólicas ou a alguns dos métodos salientados na bibliografia da especialidade (*Gensh & Shaman, 1980; Henry & Rinne, 1984*).

Devemos considerar os valores de *shares* de audiências dos canais generalistas e agregar os restantes num factor residual, uma vez que os canais por cabo apresentam valores reduzidos para os períodos em análise. No entanto, *a priori* as previsões de valores de *ratings* devem ser realizadas recorrendo a outros indicadores que apresentem uma maior estabilidade e que se irá avaliar no Apêndice F.6 Os autores Henry e Rinne (1984) recorrem a análises Logit e a algumas variáveis explicativas. Salientam que a Regressão Linear não corresponde à

metodologia mais adequada neste contexto e que os seguintes factores são os mais relevantes para tentar prever os valores de *shares* de audiências:

- A preferência por determinado programa quantificada como o número de pessoas que atribuem a classificação de bom ou muito bom ao evento. Esta variável pode ser quantificada com base nos índices de atractividade ou adesão. Atendendo às características da exposição televisiva e que constam da segunda questão em avaliação nesta tese (*ver capítulo V*) provavelmente este factor não deverá apresentar uma grande capacidade de previsão dos dados;
- Os valores de *shares* de audiências dos programas adjacentes. Nas análises a realizar aos *ratings* também observamos alguma correlação entre os valores das observações adjacentes, com uma maior associação face aos valores de lead out. Contudo, não sendo os seus valores conhecidos também não constitui uma variável com capacidade de previsão;
- O número de anos que o programa consta da grelha de programação. Representará o grau de familiaridade com determinado programa e não as preferências dos telespectadores;
- Se o programa alterou a sua posição na grelha face à temporada anterior. Salientam que a alteração da posição de um programa condiciona os resultados. Com base nas análises realizadas verificamos que o suporte temporal corresponde a um dos principais factores que condiciona a variabilidade dos indicadores de audimetria.

- A tipologia programática do programa e as sinergias entre as tipologias face aos programas adjacentes. Verificaram que as sinergias não são relevantes e que apenas duas tipologias apresentam alguma repercussão. Por exemplo, os programas transmitidos em séries apresentam maiores valores de *shares*;

- O canal de transmissão, caso exista por parte do painel a predisposição para se desenvolverem comportamentos de lealdade ao canal. Verificam que a repercussão do canal nos valores de *shares* envolve um grande horizonte temporal e que existe alguma predisposição para se desenvolverem comportamentos de lealdade.

A imagem das empresas corresponde ao seu modelo de negócio e ao centro da sua orientação estratégica. A imagem de um canal corresponde à percepção que os telespectadores apresentam da estação televisiva e o seu posicionamento à percepção comparativa do público entre os diversos canais.

Os canais generalistas em geral posicionam-se de forma muito semelhante, uma vez que todos pretendem maximizar as suas audiências. A diferença entre os canais públicos e privados torna-se subtil, como acontece com a RTP1 e com os canais privados.

Os canais públicos preferem uma imagem mais aberta e que transmita uma imagem universal e interpele a totalidade da população sem excluir nenhum grupo, o que na realidade acaba por não se verificar, existindo segundo *Fernandes (2001)* um perfil diferenciado entre os diversos canais e que se confirmará no *Apêndice F.5*.

- A duração do programa que poderá ser meia hora ou uma hora, uma vez que uma maior duração pode levar a menores valores de *shares*. Segundo os autores esta variável não influencia os resultados.

F.2. ANÁLISES EXPLORATÓRIAS DE PREVISÃO DOS INDICADORES DE AUDIMETRIA QUE CONDICIONAM A CONSTRUÇÃO DOS PLANOS DE *MEDIA*

As análises de previsão de determinados indicadores de audimetria são determinantes para o sector televisivo. O determinismo estrito defende que o estado presente de um sistema, e, por vezes, passado fixa de maneira unívoca o seu estado em qualquer instante futuro. Segundo a Mecânica Clássica, o determinismo estrito, também, é válido para qualquer sistema completamente isolado do resto do universo, ou seja, para um sistema cuja evolução não é afectada de maneira significativa pelo ambiente. Consequentemente, por circunstâncias que, certamente, não se verificam na área da audimetria, onde estão em análise comportamentos de pessoas continuamente sujeitos a uma diversidade de estímulos externos e internos e variáveis no decorrer do tempo, o que coloca grandes incógnitas relativamente à viabilidade das análises de previsão que apresentem um grau de precisão que possa ser considerado satisfatório. Inclusive em sistemas deterministas a previsão “determinista” poderá não ser viável, pois podemos não ter acesso à totalidade dos factores que influenciam a sua evolução em determinado momento.

O termo “determinismo” é uma designação ontológica, pois refere-se à natureza do mundo, ao passo que “previsibilidade” é um termo epistemológico, relativo à capacidade que temos de conhecer o futuro e que depende do conhecimento que detemos a respeito do funcionamento dos sistemas no decorrer do tempo. Apesar da imprevisibilidade para resultados de medições individuais e que, também, se verifica maioritariamente na área da audimetria, a Física Quântica permite que sejam realizadas previsões relativamente precisas sobre as frequências estatísticas com as quais diferentes resultados ocorrem no decorrer do tempo. Quer dizer, é observado um determinismo estatístico que, também, poderá ser relevante nesta área, estando a previsão dos indicadores de audimetria condicionada pela sua evolução agregada, isto é, referente a um conjunto de indivíduos. Se um sistema não for estritamente determinista, tem-se uma situação de indeterminismo, probabilismo, “aleatório”, “caótico” ou de “estocasticidade” e de “perda de causalidade”. Por conseguinte, devemos avaliar a viabilidade de previsão dos indicadores que condicionam o investimento publicitário.

F.2.1. Análises exploratórias de previsão das taxas médias audiências para o total da TV

F.2.1.1. Introdução e objectivos

Neste capítulo constam algumas análises exploratórias que visam a avaliação da viabilidade da previsão do indicador que quantifica a exposição global ao *media*. Os principais objectivos das análises a realizar correspondem aos seguintes:

(1) Segmentação da emissão diária atendendo à dimensão e variabilidade da variável em estudo.

Estas análises são relevantes, uma vez que podem condicionar as análises explicativas e de previsão multivariadas a serem realizadas posteriormente. A segmentação das séries temporais viabiliza a codificação dos dados em representações mais adequadas à implementação de determinadas metodologias de *Data Mining* proposicional e Temporal. A segmentação dos dados poderá permitir, também, indirectamente realizar a previsão deste indicador, através do *clustering* dos segmentos horários identificados.

Os factores temporais influenciam a composição sociodemográfica da audiência, a estruturação da grelha de programas e as estratégias de programação e de contra-programação adoptadas. Pretendemos verificar qual é a dimensão dos horizontes temporais que apresentam características relativamente estáveis quanto à evolução desta variável, assim como, também, identificar períodos horários onde poderá ser mais facilmente previsível por apresentar uma menor variabilidade. Uma vez identificados os segmentos horários, talvez se possa prever esta variável considerando os protótipos que os representam.

A segmentação das séries temporais permite a sua transformação em séries nominais que viabilizam a posterior aplicação de metodologias de *Data Mining* que inter-relacionam e caracterizam a evolução das séries temporais multivariadas dos principais indicadores de audimetria.

(2) Avaliar a viabilidade de previsão da variável *Tvr_Período%*. Para o desenvolvimento de modelos explicativos da evolução das audiências programáticas será necessário considerar a variável *Tvr_Prog%*, a qual poderá ser calculada em função dos períodos horários que determinado programa abrange.

Consequentemente a previsão de *Tvr_Período%* poderá ser mais relevante, quando se pretende contribuir para a previsão dos valores de *ratings* e neste apêndice recorreremos à análise das suas séries temporais. Pretende-se a análise de uma grelha de programação que exclua a transmissão de eventos desportivos e de outras tipologias programáticas que apresentam uma maior repercussão na evolução desta variável, gerando evoluções atípicas que devem ser alvo de análises diferenciadas.

F.2.1.2. Dados e metodologias

As bases de dados correspondem às que, também, foram utilizadas e descritas no capítulo anterior. As metodologias a utilizar para a realização das análises propostas procedem da Análise Simbólica ou do *Data Mining* Proposicional e/ ou Temporal.

(I) Segmentação da emissão diária atendendo aos valores de *Tvr_Período%* para o universo televisivo.

Um dos objectivos deste apêndice traduz-se na identificação de períodos horários com evoluções comparáveis da variável em estudo e que contribuam para a sua posterior previsão. Determinados segmentos horários podem apresentar uma maior estabilidade de valores. O *Data Mining* Temporal aborda a área da segmentação das séries temporais, desenvolvendo algoritmos de agregação de segmentos com base no seu grau de similitude, entre os quais, se destaca o “Sliding Window”, em que um segmento será definido até exceder uma componente de erro. Nos algoritmos “Top-Down” a série temporal será recursivamente fraccionada até se satisfazer um critério de paragem e nos algoritmos “Bottom-Up” os segmentos são agregados até uma condição de paragem ser satisfeita.

No entanto, nestas análises uma vez que está em causa a identificação de períodos comparáveis ao nível da variabilidade e evolução desta variável será mais relevante a utilização de metodologias simbólicas. Pretendemos segmentar a emissão diária em períodos horários mais homogêneos que possam orientar as futuras análises, reduzindo-se, assim, se necessário a dimensão temporal a uma variável categórica nas análises de Regressão não linear.

A classificação horária adoptada pela indústria televisiva, poderá não ser a mais adequada e representativa dos dados em análise. Com base nas análises de segmentação, talvez seja possível em paralelo realizar a sua previsão, recorrendo ao *clustering* dos segmentos horários previamente identificados. Os algoritmos de DM não simbólicos realizam uma segmentação que não necessariamente será a mesma para cada dia considerado, dificultando a identificação de faixas horárias comparáveis. Associando um intervalo de variabilidade de *Tvr_Período%* a cada um dos períodos com uma duração de meia hora e considerando a totalidade da emissão diária é possível a sua segmentação recorrendo a metodologias simbólicas.

Do ponto de vista do DM simbólico vamos recorrer às Redes SOM simbólicas e na área Simbólica a alguns algoritmos de *clustering*, como, por exemplo, aos algoritmos de *clustering* divisivo e a medidas de dissemelhanças para objectos simbólicos booleanos. Com base nas análises exploratórias realizadas, a evolução diária desta variável é semelhante dia após dia, consequentemente poderá ser viável a agregação dos seus valores.

A segmentação será realizada atendendo aos dados referentes à totalidade do painel. A segmentação das séries temporais numéricas originais, também, nos permite obter uma representação simbólica que viabiliza o seu *clustering* recorrendo a *templates* e a aplicação de inúmeras outras metodologias, como, por exemplo, o *matching* de ST.

(II) Previsão de Valores de *Tvr_Período%*

O TSDM inclui diversas áreas abrangendo desde o “Rule Discovery” até ao “Novelty Detection”. A previsão de dados temporais, decorre essencialmente dos resultados

proporcionados por metodologias que não são da área previsão. Do ponto de vista da previsão propriamente dita, os procedimentos paramétricos de análise de séries temporais incluem modelos lineares e não lineares, univariados ou multivariados, bayesianos ou não bayesianos e de previsões pontuais ou intervalares. Para além dos métodos tradicionais de análise de séries temporais, como, por exemplo, Arima, Médias Móveis, *Holt Winters*, *Smoothing Exponencial*, *Box-Jenkins*, entre outros, novos algoritmos têm sido desenvolvidos.

Uma vez que o período em análise corresponde a uma período característico a nível médio do consumo televisivo, não existe nestas séries temporais tendência, nem sazonalidade, mas, sim, uma periodicidade diária, inviabilizando a aplicação da maioria das metodologias clássicas de análise de ST. Em futuras análises a realizar considerando um horizonte temporal mais abrangente podemos recorrer a metodologias mais clássicas de análise de séries temporais, assim como ao TSDM que visa a análise de grandes bases de dados por forma a identificar regularidades que permitam um melhor conhecimento do processo subjacente. O TSDM abrange, por exemplo, a representação e modelação das sequências de dados, a definição de medidas de semelhança entre sequências e por fim a classificação ou *clustering*.

Após as análises de caracterização desta variável, uma vez que, por vezes, também se verifica alguma estabilidade na sua evolução, também se irá avaliar a viabilidade de recorrer a modelos de análise de séries temporais de valores de *Tvr_Período%*. Para o efeito utilizar-se-á algoritmos de Regressão não paramétrica, como, por exemplo, a Regressão *K-Nearest Neighbour* da área do *Instance Based Learners* (IBL), a Regressão Kernel, a Regressão Polinomial, a Regressão Local Linear e a Regressão Local Polinomial.

Os dados em análise correspondem a uma série temporal de valores discretos. No decorrer da emissão diária esta variável apresenta dois valores máximos de consumo televisivo que alternam com períodos de evolução linear ou polinomial crescente ou decrescente. Não se observa uma evolução globalmente linear ou polinomial nesta variável, mas poderá sê-lo localmente. A utilização de modelos locais apresenta algumas mais-valias relativamente ao ajuste de modelos paramétricos em subconjuntos de dados que apresentem uma evolução linear. Também, podemos recorrer às Redes Neurais e a Árvores de Regressão.

F.2.1.3. Resultados

Neste capítulo consta a avaliação da viabilidade de recorrer a algoritmos de segmentação, *clustering* e previsão de séries temporais e a metodologias simbólicas para prever a variável *Tvr_Período%*.

I. Previsão de Valores de *Tvr_Período%*

Face aos resultados alcançados, em que verificámos uma significativa proximidade nas tendências médias deste indicador no decorrer dos anos em análise, parece viável a construção de modelos univariados de previsão de *Tvr_Período%*, sendo posteriormente necessário a correcção destes valores em determinados períodos essencialmente associados à transmissão de jogos de futebol. Analisando as séries temporais dos seus valores históricos podemos proceder ao cálculo destes coeficientes.

Nesta tese não se pretende uma análise exaustiva a este nível, sendo mais relevante a identificação de interacções entre as variáveis de audimetria com repercussão ao nível da evolução das audiências. Consequentemente, apenas para não deixar uma lacuna a este nível, para a análise das séries temporais de valores de *Tvr_Período%* vamos proceder à aplicação de algumas metodologias de Regressão não Linear de *Data Mining*. Avaliou-se a possibilidade de aplicação de modelos *K-Nearest Neighbour*, a Regressão Kernel, a Regressão Polinomial, a Regressão Local Linear e Local Polinomial e as Redes Neurais MLP para a previsão de valores da variável *Tvr_Período%* ou de *Tvr_Prog%*.

Pretende-se a identificação de um modelo que se ajuste o melhor possível aos dados históricos e diferenciam-se as análises em função do horizonte temporal em análise, isto é, dados semanais e de fim-de-semana.

Ia. Resultados das análises de MBR (Regressão não paramétrica)

Na apresentação dos resultados vamos privilegiar o mesmo horizonte temporal que consta das análises de caracterização da evolução dos principais indicadores de audimetria, atendendo aos motivos anteriormente referidos.

Da aplicação destes algoritmos conseguimos no máximo uma variância explicada de 70%, 75,7%, 75% para os dados semanais e para os anos de 2004, 2003 e 2002 respectivamente.

A variância explicada calculou-se através da expressão representada por R^2_{test} , isto é:

$$R^2_{test} = 1 - \frac{MSE_{test}(\hat{y})}{\frac{1}{n_{test}} \sum_{i=1}^{n_{test}} (y_i^{test} - \bar{y}^{test})^2}$$

onde

$$MSE(\hat{y}) = \frac{1}{n_{test}} \sum_{i=1}^{n_{test}} (y_i^{test} - \hat{y}_i^{test})^2$$

Atendendo aos dados de fim-de-semana as percentagens são inferiores, apresentando os valores de 58,7%, 56,7%, 46,9% para os mesmos anos.

As percentagens de variância explicada refere-se a amostras de teste que correspondem a 30% das observações totais.

A expressão ECMP representa o erro quadrático médio de previsão de uma amostra de teste onde

$$(x_i^{test}, y_i^{test}), i = 1, \dots, n_{test}$$

corresponde à amostra de teste e $\hat{r}_h$ ao estimador não paramétrico construído com um parâmetro de alisamento h , utilizando a amostra de treino que minimize esta grandeza.

$$ECMP_{test}(h) = \frac{1}{n_T} \sum_{i=1}^{n_{test}} (Y_i^{test} - \hat{r}_h(x_i^{test}))^2$$

O erro quadrático médio relativo (RSE) define-se pela seguinte expressão:

$$RSE = \frac{1}{n} \sum_{x \in D} \frac{(y(x) - \hat{f}(x))^2}{(y(x) - \bar{f})^2}$$

onde $\bar{f} = \frac{1}{n} \sum_{x \in D} f(x)$

O numerador de RSE corresponde ao ECMP. A Regressão não paramétrica pretende identificar a função $y_i = r(x_i) + \varepsilon_i$ apresentando vantagens face ao ajustamento de modelos lineares ou polinomiais em fracções do domínio da função. As observações são ponderadas em função da sua distância relativamente à observação a prever. O peso de uma observação em $r(t)$ será calculado através da seguinte expressão:

$$w_i = w(t, x_i) = \frac{K(\frac{x_i - t}{h})}{\sum_{j=1}^n K(\frac{x_j - t}{h})}$$

onde h corresponde a um parâmetro de escala que controla a concentração do peso total na vizinhança de t .

Posteriormente, para uma Regressão linear local resolve-se o seguinte problema de mínimos quadrados ponderados:

$$\min_{a,b} \sum_{i=1}^n w_i (y_i - (a + b(x_i - t)))^2$$

onde os parâmetro a e b dependem de t .

A recta de regressão ajustada localmente na vizinhança de t será $L_t(x) = a(t) + b(t)(x - t)$.

Como função K privilegia-se a Gaussiana representada por $(1 / \sqrt{2\pi}) \exp(-x^2 / 2)$.

Para estimar um polinómio de grau q calculam-se os pesos como no contexto linear e posteriormente recorre-se à seguinte expressão:

$$\min_{\beta_0, \dots, \beta_q} \sum_{i=1}^n w_i (y_i - (\beta_0 + \beta_1(x_i - t) + \dots + \beta_q(x_i - t)^q))^2$$

□ Resultados das análises realizadas com os dados semanais

Considerando os dados referentes aos dias laborais de 2004, as metodologias que maximizam a variância explicada correspondem a Regressões Locais e à Regressão Kernel.

➤ Regressão ultra local, com uma função de ponderação Gaussiana e uma largura de Kernel no valor de 0,015. Corresponde a uma Regressão local linear com a seguinte equação: $y = 1 + (\text{hora} - 5) / 20$ e que representa 70% de variância explicada sobre amostra de teste. Consideramos como métrica de distância a euclidiana $D(x, x') = \sqrt{(x - x')^T \Sigma (x - x')}$ e uma função de ponderação Gaussiana representada por $w_k = \exp(-D(x_k, x_{query})^2 / K_w)$;

Onde K_w representa a largura de Kernel e posteriormente calcula-se:

$$\hat{\beta} = \arg \min_{\beta} \sum w_k^2 (y_k - \beta^T x_{query})^2$$

$$y_{previsto} = \beta^T x_{query}$$

O Gráfico F.2.1.3.I.1 representa as previsões de valores de $Tvr\_Período\%$ e o respectivo intervalo de confiança a 95%, correspondendo o eixo das abcissas aos intervalos de meia hora considerados para o cálculo dos valores de $Tvr\_Período\%$. O Gráfico F.2.1.3.I.2 representa os valores dos resíduos e, também, o respectivo intervalo de confiança a 95%.

➤ Regressão Kernel com 6KNN adoptando uma função de ponderação Gaussiana e uma largura de Kernel no valor de 0.015625 e que apresenta 70% de variância explicada (Gráfico F.2.1.3.I.3 e Gráfico F.2.1.3.I.4).

Neste caso o valor da previsão consiste na média ponderada dos seis valores de $Tvr\_Período\%$ mais próximos a uma observação e com um factor unitário e para os restantes três segundo a seguinte função de ponderação:

$$\hat{y} = \frac{\sum w_i y_i}{\sum w_i} \text{ e } w_i = \exp\left(\frac{-D(x_i, query) H^2}{K_w^2}\right)$$

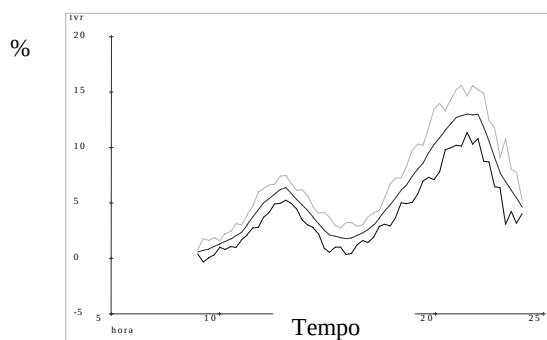


Gráfico F.2.1.3.I.1 - Previsão com uma Regressão Ultra-Local Linear

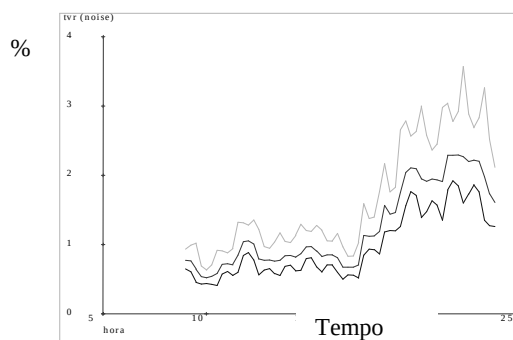


Gráfico F.2.1.3.I.2 - Representação dos valores de resíduos da Regressão Local Linear

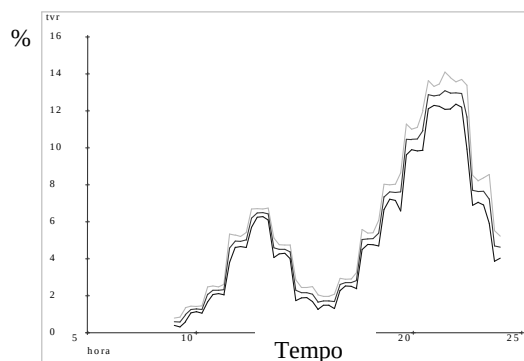


Gráfico F.2.1.3.I.3 - Previsão com uma Regressão Kernel com 6 KNN

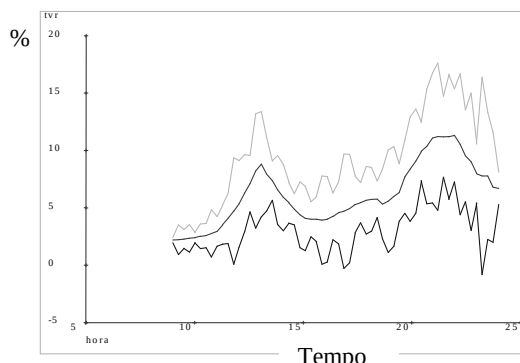


Gráfico F.2.1.3.I.4 - Representação dos resíduos da Regressão Kernel 6KNN

Sem dúvida, que os programas e os períodos directamente adjacentes a determinado programa são os que mais se aproximam quanto aos valores destes índices, representando os 6 KNN um horizonte temporal de aproximadamente 3 horas. Este último modelo apresenta como inconveniente, as discontinuidades existentes nos valores previstos e o erro é uma vez mais superior a partir das 20H.

□ Resultados das análises realizadas com os dados de fim-de-semana

O melhor modelo considerando os dados de fim-de-semana de 2004 corresponde a uma Regressão Quadrática ultra local, com uma função de ponderação Gaussiana e uma largura de kernel de 0.0156 segundo a função: $y=1+(hora-5)/20+(hora-5)/20)^2$. Apresenta 58,7% de variância explicada.

Ao fim-de-semana a variabilidade nos dados é superior no decorrer de todo o dia e o IC para os valores das previsões e para os resíduos será superior nos períodos de maiores audiências.

Na Regressão Local os coeficientes de regressão são dados pela seguinte expressão:

$$\hat{\beta} = ((WX)^T (WX))^{-1} ((WX)^T (WY))$$

onde X representa a matriz que tem uma coluna para cada variável independente e Y uma matriz com uma coluna para cada variável dependente, $\hat{\beta}$ são os coeficientes de regressão e W a matriz diagonal com os factores de ponderação das observações.

Na Regressão Local Quadrática adicionamos termos quadráticos à matriz X, por exemplo, $1, x_1, x_2, x_1^2, x_1 x_2, x_2^2$.

□ Alguns resultados para o ano 2003

O melhor modelo de Regressão para os dados semanais de 2003 corresponde a uma Regressão Linear local com os mesmos termos que foram identificados para o ano de 2004 (R^2_{test} com o valor de 75,7%). Utiliza também a função Gaussiana e uma largura de kernel com o valor de 0.03125. Para os dados de fim-de-semana o melhor modelo corresponde ao mesmo modelo identificado para o período semanal (R^2_{test} com o valor de 56,7%). O intervalo de confiança dos valores previstos para ambos os períodos é menor relativamente aos do ano de 2004.

□ Alguns resultados para o ano 2002

O melhor modelo de Regressão para os dados semanais de 2002 coincide com o identificado para 2004, apresentando um R^2_{test} de 75%. Corresponde a uma Regressão Ultra Local de mínimos quadrados, com uma função de ponderação Gaussiana e uma largura kernel de 0.015625 e com a seguinte equação: $y=1+(hora-5)/20+(hora-5)/20 \times (hora-5)/20$. Para os dados de fim-de-semana o melhor modelo corresponde a uma Regressão Kernel, com uma função de ponderação Gaussiana e uma largura de kernel de 0.03125 (R^2_{test} de 46,9%). Com os modelos de Regressão Local obtemos razoáveis percentagens de variância explicada para os dias laborais, com valores compreendidos entre os 70% e os 75,7%, sendo esta percentagem inferior para o fim-de-semana com os valores compreendidos entre os 46,7% e os 58,7%. Os melhores resultados para os cinco anos de maior instabilidade no desempenho dos canais correspondem aos anos de 2003 e 2004.

II. Resultados que decorrem da aplicação das Redes Neurais e das metodologias de Médias Móveis

Para a implementação de uma Rede Neuronal MLP de previsão de valores de *Tvr_Prog%* recorreu-se como variáveis de entrada às variáveis nominais referentes ao canal, se o programa é transmitido durante a semana ou ao fim-de-semana e ao período horário. Estas variáveis foram codificadas com as seguintes variáveis dummy: duas para a variável canal que apresenta os níveis SIC, TVI e RTP1, uma variável *dummy* para a variável dicotómica semana ou fim-de-semana e cinco para o período horário que resulta das análises de segmentação simbólicas realizadas. Como funções de activação (f_i) utilizou-se a *Sigmoid* (1) e a *Tangente Hiperbólica* (2), isto é:

$$y = f(x) = \frac{1}{1 + e^{-x}} \quad (1) \quad \text{e} \quad \tanh(x) = \frac{e^{2x} - 1}{e^{2x} + 1} \quad (2)$$

A saída de cada neurona de entrada será dada Por:

$$y_i(t) = f_i(\sum_j w_{ij} x_j - \theta_i)$$

A saída da Rede z_k^μ corresponde a:

$$z_k^\mu = \sum_j w_{kj} y_j^\mu - \theta_k = \sum_j w_{kj} f(\sum_i w_{ji} x_i^\mu - \theta_j) - \theta_k$$

O algoritmo pretende minimizar $E(w_{ij}, \theta_j, w_{kj}, \theta_k)$ dado por:

$$(1/2) \sum_\mu \sum_k \left[t_k^\mu - f(\sum_j w_{kj} y_j^\mu - \theta_k) \right]^2$$

A melhor Rede MLP, por exemplo, atendendo aos dados de 2004 segundo o algoritmo gradiente descendente é da tipologia 7-4-1, apresentando apenas uma capa oculta.

Utilizou-se a validação cruzada com a amostra inicial subdividida em amostra de teste (25% das observações), treino (55% das observações) e de validação (20% das observações). Apresenta um valor de MSE de 6.4 para a amostra de teste e de 6.25 para a amostra treino. Discretizando a variável de saída em intervalos de 3% de amplitude, segundo a matriz de classificações apenas 57,6 % das observações foram correctamente classificadas para a amostra de teste. Portanto estas variáveis não conseguem prever correctamente a variável *Tvr_Prog%* por aplicação de uma Rede Neuronal MLP. Os restantes anos, também, não apresentam resultados satisfatórios, recorrendo a esta metodologia. Atendendo aos dados de 2004 um ajustamento de médias móveis com N=3 conduz a valores de MSE inferiores aos das Redes MLP (MSE de 2,8 e com N=2 um MSE no valor de 1,5).

Consequentemente, considerando as Regressões não lineares os resultados não são no geral muito satisfatórios e devem ser equacionadas alternativas. Contudo, nos resultados da bibliografia internacional ao nível da previsão dos valores de *ratings*, por vezes, consideram-se valores de 70% de variância explicada como “bons” atendendo às dificuldades que estas análises colocam ao sector

III. Análises de Segmentação de *Tvr_Periodo%*

Foram realizadas análises recorrendo à base de dados global (anos 2000-2008) e para cada um dos anos individualmente por aplicação de algumas metodologias simbólicas.

Nomeadamente recorremos aos algoritmos de *Clustering* Divisivo, às Redes Neurais SOM simbólicas, assim como a medidas de dissimilaridade entre objectos simbólicos. A base de dados simbólica (*Quadro F.2.1.3.III.1*) corresponde aos conceitos que traduzem os intervalos de meia hora ao longo da totalidade da emissão diária caracterizados em função dos intervalos de variabilidade de *Tvr_Período%* para os vários anos considerados. Por conseguinte, correspondem a objectos simbólicos booleanos.

	2000	...	2008
10H-10.30H	$[\alpha_{i1}, \beta_{j1}]$	...	$[\alpha_{i1}, \beta_{j5}]$
...	...	...	...
...	...	...	...
24H-24.30H	$[\alpha_{im}, \beta_{j1}]$	...	$[\alpha_{im}, \beta_{j5}]$

Quadro F.2.1.3.III.1 – Base de dados simbólica para a análise de *Tvr_Período%*

Excluíram-se as observações que correspondem a programas que geram observações atípicas, por forma, a caracterizar as tendências médias existentes nos dados e não as pontuais. As situações atípicas devem ser alvo de análises diferenciadas.

Uma vez identificados os segmentos horários podemos realizar o *clustering* de ST para identificar tendências médias em cada um desses períodos recorrendo a metodologias de *Data Mining* Temporal. As análises de segmentação recorrendo a metodologias simbólicas talvez nos permitam calcular intervalos de previsão.

As variáveis simbólicas são da tipologia intervalo ($x_{uj} = [\alpha_{uj}, \beta_{uj}]$ e $x_{vj} = [\alpha_{vj}, \beta_{vj}]$) e como medidas de dissimilaridades entre n objectos simbólicos descritos por p variáveis da tipologia intervalo podemos recorrer às seguintes:

Hausdorff representado pela expressão:

$$\delta_j(x_{uj}, x_{vj}) = \max \left\{ |\alpha_{uj} - \alpha_{vj}|, |\beta_{uj} - \beta_{vj}| \right\}$$

L1 representado pela expressão:

$$\delta_j(x_{uj}, x_{vj}) = |\alpha_{uj} - \alpha_{vj}| - |\beta_{uj} - \beta_{vj}|$$

L2 representado pela expressão :

$$\delta_j(x_{uj}, x_{vj}) = (\alpha_{uj} - \alpha_{vj})^2 - (\beta_{uj} - \beta_{vj})^2$$

Posteriormente, combinam-se os p índices de dissimilaridade para obter uma medida global, sendo esta dada por:

$$D: ExE \rightarrow R^+ : (x_u, x_v) \rightarrow d(x_u, x_v) = \sum_{j=1}^p \delta_j^2(x_{uj}, x_{vj})^{1/2}$$

IIIa. Resultados das análises realizadas com a totalidade da amostra (2000-2008)

Da aplicação das Redes Neurais SOM simbólicas não conseguimos resultados satisfatórios, com a sobreposição dos protótipos dos segmentos horários e uma forte desigualdade no número de observações incluídas em cada conjunto. A maioria das observações inclui-se no segmento associado à maior variabilidade de valores para todos os canais.

Por aplicação de outros algoritmos de *clustering* simbólicos, também não foi viável proceder à segmentação dos dados recorrendo à globalidade dos dados. Alternativamente calcularam-se medidas de dissimilaridade atendendo aos valores de variabilidade de *Tvr_Período%*. Calcularam-se as medidas de dissimilaridade entre os vários objectos simbólicos para verificarmos quais são os períodos que mais se aproximam entre si.

Uma vez que as variáveis em análise são da tipologia intervalo, os objectos simbólicos correspondem ao tipo booleano. Os resultados traduzem a sequencialidade existente nos dados e os padrões diários identificados, existindo uma evolução muito gradual na evolução de *Tvr_Período%*, o que torna difícil a constituição de segmentos bem definidos. As soluções proporcionadas, por vezes apresentam alguma discrepância, tendo-se seleccionado os resultados que apresentam um maior número de faixas horárias adjacentes e que são mais lógicos atendendo às análises exploratórias realizadas.

Por exemplo, segundo a aplicação da medida SO1 à totalidade dos dados, são mais semelhantes os períodos horários que decorrem entre as 20H e as 23.30H durante a semana e entre as 20H e as 23H ao fim-de-semana (Fig. F.2.1.3.III.1).

A medida SO1 que traduz a semelhança entre $d_i(A_j, B_j)$ representa-se por

$$d_q^i(a, b) = \sqrt[q]{\sum_{j=1}^p [w_j d_i(A_j, B_j)]^q}$$

e apresenta valores no intervalo $[0,1]$.

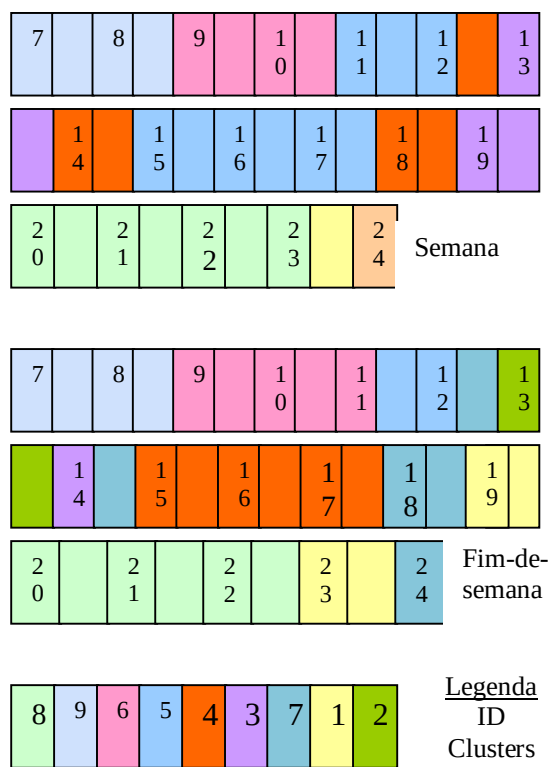


Fig. F.2.1.3.III.1 – Segmentos horários que resultam da aplicação da medida de dissimilaridade simbólica SO1

Os restantes conjuntos de períodos horários estão constituídos pelos seguintes conjuntos:

(Seg. 1) Ao fim-de-semana, o pré-prime-time equipara-se ao final da noite, ao nível da variabilidade de valores. O final da noite semanal apresenta uma evolução semelhante ao período equivalente do fim-de-semana, mas o mesmo não sucede com o programa que antecede o espaço informativo da noite;

(Seg. 2) O espaço informativo da tarde ao fim-de-semana diferencia-se do semanal e não é equiparável a mais nenhum período;

(Seg. 3) A hora do almoço semanal equipara-se ao pré-prime-time e aos períodos sucessores do espaço informativo da tarde ao fim-de-semana;

(Seg. 5) A tarde do período semanal pode ser equiparada ao final da manhã ao longo de toda a semana;

(Seg. 6) O período matinal equipara-se no decorrer de toda a semana;

(Seg. 8) A exposição televisiva maximiza-se entre as 20H e as 23H no decorrer de toda a semana;

(Seg. 7 e 4) Os segmentos 7 e 4 incluem vários períodos não adjacentes, mais difíceis de justificar.

Destes resultados verificamos que existe uma faixa horária entre as 20H e as 23H que é equiparável e na qual se maximiza a exposição televisiva do painel.

Ressalta, também, a disparidade da evolução deste indicador no decorrer dos dias laborais e ao fim-de-semana. Não existe sequencialidade nos períodos que constituem os diferentes conjuntos, abrangendo, por vezes, unidades temporalmente distantes. Nem sempre, uma maior proximidade torna os períodos horários mais semelhantes. Para a previsão das audiências deve-se ter em atenção as faixas horárias, as quais apresentam um diferenciado grau de variabilidade.

Para facilitar a interpretação dos resultados aplicaram-se as seguintes medidas de dissimilaridade apenas aos dados de fim-de-semana: U1, SO1, U2, S6, SO4. Recorreu-se também à medida de Ichino-Yaguchi dada por

$$\varphi_j(A, B) = \frac{|A_j \oplus B_j| - |A_j \otimes B_j| + \gamma(2|A_j \otimes B_j| - |A_j| - |B_j|)}{|A_j| + |B_j|}$$

Onde $|A_j \oplus B_j|$ representa a união de dois intervalos e $|A_j \otimes B_j|$ a intersecção de dois intervalos.

Os melhores resultados decorrem da aplicação da medida SO4 representada por

$$SO4 = d_2(a, b) = [\Pi(d_a \oplus d_b) - \Pi(d_a \otimes d_b) + \gamma(2\Pi(d_a \otimes d_b) - \Pi(d_a) - \Pi(d_b))] / \Pi(d_a^E)$$

Onde

$$\Pi(a^E) = [Y_1 \in Y_1] \wedge \dots \wedge [Y_p \in Y_p]$$

De acordo com o mapa de dissimilaridade calculado com base na medida SO4, entre as 20.30H e as 23H existe um período que corresponde às maiores audiências. Seguem-se os períodos que correspondem a pontos de transição nas audiências por ordem decrescente de ratings: 20H-20.30H com alguma separação face ao período de maiores audiências, 23H-23.30H que também se distancia do período que lhe segue: 19.30H-20H, 23.30H-24H, 13H-14H. Posteriormente, existe um conjunto mais indiferenciado de períodos que se diferencia das 12H-12.30H e este por sua vez do matinal (Fig. F.2.1.3.III.2).

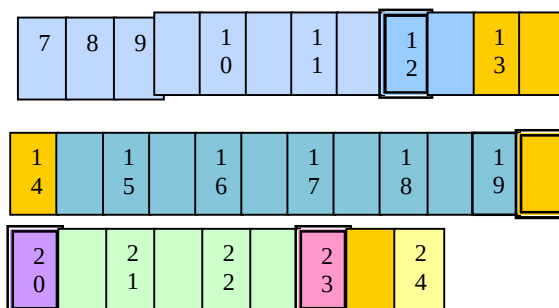


Fig. F.2.1.3.III.2 – Segmentos horários para os dados de fim-de-semana recorrendo à medida de dissimilaridade simbólica SO4

Para os dias laborais e por aplicação da mesma medida SO4 existe uma maior separação entre os períodos de maiores audiências o que denota uma evolução mais acentuada dos valores de $Tvr\_Período\%$ (Fig. F.2.1.3.III.3).

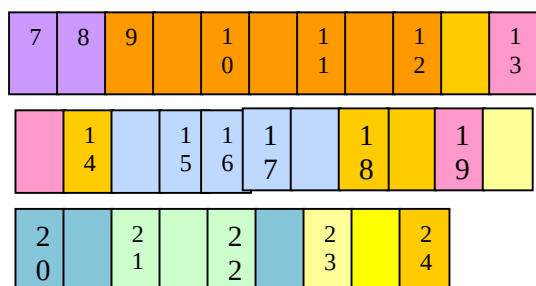


Fig. F.2.1.3.III.3 – Segmentos Horários para o período semanal recorrendo à medida SO4

As maiores audiências associam-se às faixas horárias 21H-22.30H, seguem-se as 20.30H-21H, 22.30H-23H, 20H-20.30H. Com uma acentuada separação face aos restantes períodos segue-se a faixa das 23H-23.30H. Posteriormente, a uma grande distância o conjunto 19.30H-20H agregado a 23H-23.30H e uma vez mais distanciados do conjunto 19H-19.30H agregado a 13H-14H. Posteriormente, seguem-se as faixas 24H-24.30H, 18.30H-19H e 14H-14.30H+12.30H-13H+18H-18.30H.

IIIb. Resultados das Análises Anuais

Aplicaram-se algumas metodologias de *clustering* simbólicas aos dados de cada um dos anos considerados nas análises, verificando se é viável chegar a uma segmentação dos dados em faixas horárias equivalentes. Por exemplo, atendendo aos dados de 2004 e recorrendo ao algoritmo de *clustering*

simbólico Divisivo, experimentaram-se diferentes valores de números de *clusters*.

Valores de $Tvr\_Período\%$	X_{ui}	s	C_s
$[\alpha_1, \beta_1]$	$X_{u1} = (\alpha_1 + \beta_1)/2$	1	$(X_{u1} + X_{u2})/2$
$[\alpha_2, \beta_2]$	$(\alpha_2 + \beta_2)/2$	2	$(X_{u2} + X_{u3})/2$
$[\alpha_3, \beta_3]$	$(\alpha_3 + \beta_3)/2$	3	$(X_{u3} + X_{u4})/2$
...	...	...	...

Quadro F.2.1.3.III.2 – Aplicação do algoritmo DIV para a identificação de segmentos horários

Em primeiro lugar calcula-se os pontos médios de cada intervalo X_{ui} com $s=1, \dots, m$ e os valores de corte C_s . Existem m valores de corte dando origem a partições $\{C_1, C_2\}$ calculando-se as variâncias intra-clusters $I(C_i)$ e a variância intra-cluster total $W(C_i) = I(C_i) + I(C_2)$. Escolhe-se a partição que maximiza o valor de $I_s = \Delta W(C_i)$ (Quadro F.2.1.3.III.2). Para uma solução a cinco *clusters* a variância explicada é de 91% e para nove *clusters* de 96.3%, o que, também, acarreta um maior sobre ajustamento aos dados e poderá invalidar a generalização dos resultados extraídos.

Com base neste algoritmo e para cinco *clusters* foi possível retirar as seguintes conclusões: (Fig. F.2.1.3.III.4 e Fig. F.2.1.3.III.5)

Os menores valores da variável $Tvr\_Período\%$ (valores inferiores aos 11% e com valores de variabilidade inferiores aos 5%), corresponde ao período matinal (7.30H-12H) no decorrer de toda a semana.

O período matinal minimiza os valores em dimensão e variabilidade, inclusive ao fim-de-semana. O segundo *cluster* com menores valores de *Tvr_Período%* abrange o período que antecede o espaço informativo da tarde (12H-13H), conjuntamente com 14H-18H em dias laborais ou com 15H-18H ao fim-de-semana. No período que decorre entre as 12H e as 13H verifica-se um maior acréscimo nos níveis de audiência que culminam com o espaço informativo da tarde. Este período apresenta valores entre os 11% e os 16,25% não se diferenciando o período semanal do fim-de-semana. As faixas horárias entre as 14H-18H (S) e entre as 15H-18H (FS) apresentam valores da mesma ordem de grandeza e grau de variabilidade, mas com uma evolução decrescente nos valores dos *ratings*. Ao fim-de-semana o espaço informativo da tarde consegue reter os maiores níveis de audiência por um maior horizonte temporal.

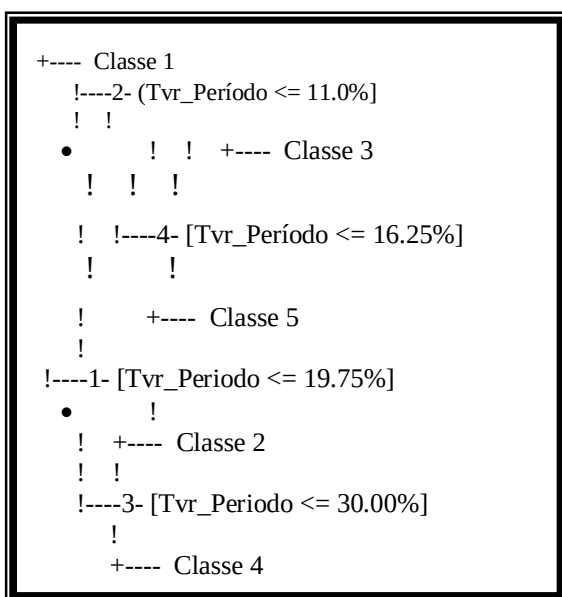


Fig. F.2.1.3.III.4 – Resultado da aplicação do algoritmo de *clustering* Divisivo a *Tvr_Período%* e para uma solução a 5 *clusters*

Os maiores valores de *Tvr_Período%* obviamente correspondem ao período que decorre desde o espaço informativo da noite até às 23H durante toda a semana. Ao nível da sua variabilidade, os valores semanais entre as 22H e as 23H são menores em relação aos valores dos restantes períodos do horário nobre e ao fim-de-semana entre as 21.30H e as 23H. Este período apresenta valores de *Tvr_Período%* superiores aos 30%. Esta segmentação da emissão diária não corresponde às faixas horárias utilizadas na prática por algumas agências de planificação de meios, que consideram horizontes temporais mais alargados e adjacentes.

O prime time, geralmente, abrange entre as 20H e as 22H e a tarde costuma ser dividida em antes e após as 16H (ao FS) ou 16.30H (S). Se utilizarmos uma solução a nove *clusters* (96,3% de variância explicada) reduzimos o período de maiores audiências durante os dias laborais a 20H-22.30H e ao fim-de-semana à classificação usada tradicionalmente (20H-22H). Os índices de audiências são superiores durante os dias laborais. Segue-se o período das 22.30H-23.30H durante toda a semana.

Repetindo as mesmas análises para o ano de 2003 e, também, por aplicação do algoritmo *DIV* e para uma solução a cinco *clusters* os valores que definem a fronteira entre classes são muito semelhantes aos identificados para 2004 e a inércia explicada é também de 91%. Tínhamos verificado que existe um forte grau de semelhança entre as evoluções de 2004 e 2003. Para o ano de 2002 e por aplicação do algoritmo *DIV* com 5 *clusters* a inércia explicada é de 90%. O ano de 2002 é o que mais se diferencia dos restantes.

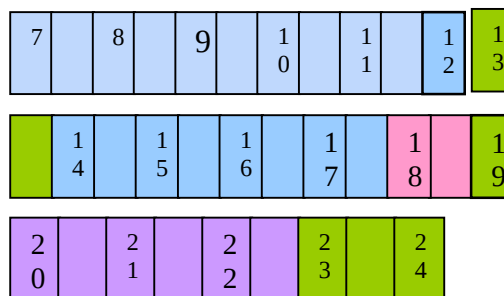
Ano 2004	Ano 2003	Ano 2002
7.30 h – 12H (S/FS)	7-12.30H (FS) 7-12H (S) 15-18H (S)	7-12H (FS) 7-10.30 H (S)
12-13 H (S/FS) 14-18H (S) 15H-18H (FS)	15-17H (FS) 12-13H (S) 14-15H (S) 18-18.30H (S)	11.30 – 12.30H (FS) 9-12.30 H (S) 14.30-17.30 H (S)
18H – 19 H (S) 18-19.30H (FS) 14-15H (FS)	12.30-13H (FS) 14-15H (FS) 17.30 – 19H (FS) 24-24.30H (S/FS)	12.30 – 20H (FS) 23.30-24.30H (FS) 12-14.30 H (S) 18-19.30 H (S) 24-24.30 H (S)
13H – 14H (S/FS) 19-20H (S) 23-24.30H (S/FS) 19.30H – 20H (FS)	13-14H (FS/S) 19-20.30 H (FS) 23-24 H (FS) 19-20H (S) 23.30-24H (S)	19.30 – 21 H (FS) 23-23.30H (FS/S) 19.30-20H (S)
20-23 H (S) 20 – 23H (FS)	20.30 – 23 H (FS) 20-23.30H (S)	21H – 22.30 H (FS) 20-24H (S)

Tabela F.2.1.3.III.1 – Descrição da constituição de segmentos horários para alguns dos anos considerados nas análises

Na Tabela F.2.1.3.III.1 consta a constituição dos *clusters*, por ordem crescente de valores de *Tvr_Período%*. Apesar da fronteira entre as classes não se diferenciar de forma significativa para os diversos anos a constituição dos *clusters* difere como consequência das alterações nos padrões de disponibilidade do painel. Por vezes, os diferenciais em termos absolutos não são muito significativos, mas atendendo à dimensão dos índices de audiência, a variabilidade relativa é consideravelmente mais significativa.

Com base na construção dos mapas de dissimilaridade verificamos que apesar de existirem alguns pontos de inflexão na evolução desta variável, sendo mais notório o seu afastamento para o período semanal, sem dúvida que as observações em períodos próximos estão bastante correlacionadas. Nas análises a realizar considera-se como solução de segmentação a que consta da seguinte Fig. F.2.1.3.III.5. Por exemplo, nas análises de Regressão não Linear consideramos esta classificação para a identificação nominal de faixas horárias, assim como para a sua conversão em séries simbólicas.

Solução de Segmentação (Dados Semanais)



Solução de Segmentação (Dados Fim-de-semana)

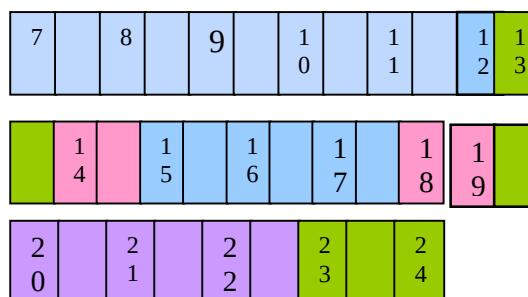


Fig. F.2.1.3.III.5 – Representação de *clusters* horários para o período semanal e de fim-de-semana

F.2.2. Análises exploratórias de previsão de valores de ratings

F.2.2.1. Objectivos

Os *ratings* representam um indicador que é essencial para a construção de planos de investimento em publicidade televisiva. Consequentemente, a sua previsão e uma maior percepção sobre os factores que condicionam a sua evolução é determinante para este sector. Neste apêndice pretende-se avaliar o grau de previsibilidade dos valores de *ratings* para períodos horários, recorrendo apenas à análise das respectivas séries temporais. As abordagens de previsão são inúmeras, desde as mais simples extrapolações ou prolongações de tendências até à análise da grelha e à procura dos cenários históricos mais semelhantes ao da observação em análise. Neste apêndice avalia-se apenas as possibilidades de análise recorrendo aos valores das suas séries temporais.

Inicialmente, através de algoritmos de *clustering* pretende-se identificar se existem tendências médias de evolução de *ratings* que sejam representativas para cada um dos canais. Posteriormente, realizam-se análises de segmentação que permitam a identificação de períodos horários mais restritos que sejam comparáveis quanto à evolução desta variável e que possam apresentar uma maior estabilidade de valores, talvez viabilizando a sua posterior previsão, por exemplo, através da aplicação de algoritmos de *clustering* de séries temporais. Em função dos resultados poder-se-á recorrer a algoritmos de previsão, essencialmente não lineares.

As abordagens metodológicas mais apropriadas para a previsão de valores de *ratings* são função dos contextos. Numa óptica explicativa avaliamos como as alterações na evolução das taxas médias de audiências de determinado canal se repercutem nos valores dos *ratings* e *shares*, não só do canal, como também, dos canais concorrentes. Pretendemos avaliar se se justifica recorrer a séries temporais multivariadas para prever este indicador e em que contextos, uma vez que no Apêndice F.6 apenas se tem em consideração a influência das variáveis estruturais de planeamento que constam das bases de audimetria.

F.2.2.2. Dados e metodologias

Uma vez que as análises propostas incidem na avaliação da viabilidade de previsão dos *ratings*, recorrendo a séries temporais univariadas, as metodologias a serem utilizadas abrangem a segmentação, *clustering* e previsão de séries temporais e descrevem-se nos seguintes parágrafos.

(1) *Clustering* de séries temporais de valores de *ratings*

As análises realizadas aos valores de *ratings* programáticos consistiram na análise gráfica das séries temporais comparando a sua evolução diária no decorrer do tempo e na aplicação de algoritmos de *clustering* simbólicos para avaliar o seu grau de estabilidade. As séries temporais por se referirem a programas apresentam diferentes comprimentos, sendo portanto inviável o cálculo de meras distâncias euclidianas entre as séries temporais originais. Neste contexto se necessário, em futuras análises, podemos adoptar metodologias de *Data Mining* Temporal que visem a representação das séries, ou seja, o seu pré processamento e a avaliação do seu grau de similitude considerando as análises inter-canais, assim como as evoluções referentes a um único canal. Será necessário recorrer a algoritmos de “Dynamic Time Wrapping” para alinhar as séries e, posteriormente, aplicar metodologias de *clustering*.

Nas análises a realizar optou-se por associar valores de *ratings* a intervalos temporais, sendo a variável designada de *Rat_Período%*, obtendo-se, assim, uma série de valores discretos e equitativamente espaçados para cada um dos canais. Após esta transformação coincidem os instantes temporais e as séries temporais passam a apresentar o mesmo comprimento, podendo-se avaliar o seu grau de similitude recorrendo ao cálculo da distância euclidiana ou ao coeficiente de correlação linear de pearson, com a posterior aplicação de algoritmos proposicionais de *clustering*. Nas análises a realizar adopta-se preferencialmente o coeficiente de correlação, uma vez que não é influenciado pela escala de medida das variáveis e pelas respectivas médias. As séries temporais apresentam variáveis que são medidas na mesma escala, uma vez que os *ratings* são percentagens e com linhas base muito semelhantes.

Clustering de ST recorrendo ao algoritmo K Média	Clustering de Séries Temporais recorrendo a Algoritmos Hierárquicos Aglomerativos
O algoritmo K Média apresenta algumas limitações, como, por exemplo, ser sensível às médias iniciais. Consequentemente, corre-se o algoritmo inúmeras vezes, por exemplo, 50 vezes com médias iniciais aleatórias e selecciona-se a solução de <i>clustering</i> que maximiza o somatório das semelhanças, considerando a totalidade dos <i>clusters</i>. O número de <i>clusters</i> mais adequado pode ser previamente identificado, recorrendo à Análise Factorial em Componentes Principais. Para reduzir a dimensionalidade das séries temporais multivariadas e identificar o número de <i>clusters</i> podemos recorrer à análise em Componentes Principais que permite a identificação dos vectores base ortogonais, não correlacionados.	Para conduzir o <i>clustering</i> das séries temporais, também, podemos recorrer a algoritmos hierárquicos aglomerativos. Pretendemos agregar as evoluções diárias em <i>clusters</i> com base no seu grau de semelhança. Cada <i>cluster</i> irá conter uma evolução diária média e, assim, provavelmente, contribuindo para a previsão deste indicador. Divide-se cada séries temporal numa sequência de M evoluções diárias. Cada evolução diária Y_j, $j=1, \dots, M$ consiste numa sequência de pares de valores (y_i, t_i), $i=1, \dots, N$, onde y_i representa o valor do <i>rating</i> e t o correspondente instante temporal. Inicia-se com M <i>clusters</i>, contendo cada <i>cluster</i> uma evolução diária. Calcula-se as diferenças entre cada <i>cluster</i> e agregam-se os dois mais semelhantes. Como resultado ficamos com M-1 <i>clusters</i> activos. Esta fase de agregação repete-se até atingirmos um único <i>cluster</i>. O resultado deste algoritmo será uma Árvore binária de 2M-1 <i>clusters</i>. Podem-se utilizar várias medidas de dissemelhança. Supondo que temos duas evoluções diárias y_i e z_i, $i=1, \dots, N$ podemos utilizar a distância geométrica, entre outras. Podemos utilizar, também, a distância euclidiana, mas esta medida apresenta a desvantagem de não permitir a utilização de diferentes linhas base e de amplitude de variação. Os <i>clusters</i> podem ser representados por protótipos que representam a média de valores ou por <i>medoids</i>. Para visualizar os resultados recorreremos a dendogramas.

Quadro F.2.2.2.1 – As metodologias de *clustering* de séries temporais de valores de *ratings* adoptadas nas análises- Fonte: Elaborado pela autora

Relativamente aos algoritmos de *clustering* de séries temporais existem vários algoritmos. Quando as séries temporais apresentam o mesmo comprimento podem ser utilizadas metodologias clássicas de *clustering* considerando cada série como um vector num espaço eucladiano. Nas análises a realizar com as séries originais no domínio temporal utilizam-se as metodologias que constam do *Quadro F.2.2.2.1*.

(2) Análises de segmentação de *ratings*: A partir dos dados simbólicos se associarmos um intervalo de variabilidade de *ratings* a cada unidade temporal é possível avaliar o seu grau de semelhança por aplicação de medidas simbólicas, como foi anteriormente implementado para a análise das taxas médias de audiências para o total da TV. Para as análises de segmentação recorreremos a metodologias de *clustering* simbólicas e conseguimos uma segmentação global dos dados que será função dos intervalos de variabilidade desta variável para o período considerado.

(3) Análises de previsão com base na análise de séries temporais de valores de *Ratings*: Para realizar a previsão de *ratings* podemos recorrer a várias perspectivas de análise destacando-se as que constam do *Quadro F.2.2.2.2*.

Clustering de Séries Temporais	Segmentação e Clustering de Séries Temporais
Através do <i>clustering</i> das séries temporais diárias que ilustram a evolução dos <i>ratings</i> e recorrendo previamente à análise em Componentes Principais.	Através da segmentação e <i>clustering</i> dos segmentos temporais, talvez seja viável calcular uma evolução média desta variável para determinados períodos horários, permitindo a sua previsão.
Algoritmos de previsão e de Regressão não paramétrica	
Ao nível da Regressão não paramétrica, os algoritmos coincidem com os adoptados para a análise das taxas médias de audiências para o total da TV. As metodologias de Regressão não paramétrica locais consideram as observações na vizinhança da observação a prever para calcular o valor da previsão. Em alternativa podemos recorrer a métodos clássicos de análise de séries temporais, como, por exemplo, as Médias Móveis ou a Regressões Trigonométricas e as Redes Neurais.	

Quadro F.2.2.2.2 – A estrutura metodológica das análises propostas para a análise dos *ratings*

F.2.2.3. Resultados

I. Análises de *Clustering* de *Rat_Período%*

Realizaram-se algumas análises de *clusters* com base na evolução diária da variável *Rat_Período%* para cada um dos canais. Como medida de dissemelhança recorreu-se ao coeficiente de correlação linear de pearson e como algoritmo de *clustering* ao K Média e a algoritmos hierárquicos. O número de *clusters* mais adequado foi definido recorrendo a análises em Componentes Principais. No gráfico F.2.2.3.I.1 constam os resultados para o canal SIC e encontram-se representados os valores médios dos *ratings* associados aos diferentes períodos horários para cada um dos *clusters* identificados.

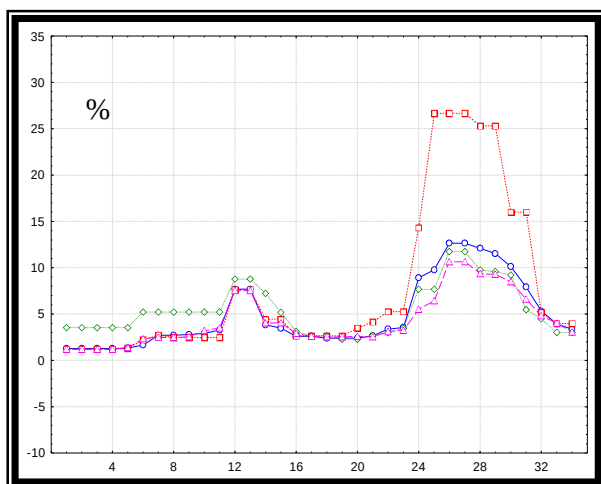


Gráfico F.2.2.3.I.1 – O *clustering* de séries temporais de valores de *ratings* para o canal SIC

Verificamos que existe uma grande proximidade de valores entre os *clusters* 2 e 4, apenas se diferenciando nos períodos de maiores audiências. O primeiro *cluster* inclui evoluções mais atípicas, enquanto que o segundo corresponde ao mais representativo da evolução dos dados.

O quarto *cluster* engloba essencialmente dias de fim-de-semana e o segundo *cluster* inclui a maioria dos dias semanais. O *cluster* 3 e 1 com um menor número de observações corresponde a dias em que foram transmitidos jogos de futebol nos canais concorrentes ou outros programas que geraram observações atípicas da variável em estudo.

Consequentemente, as evoluções diárias face a uma grelha de programação já estabelecida são próximas, existindo uma nítida diferenciação entre os dias laborais e os dias de fim-de-semana. Os eventos desportivos são responsáveis por uma evolução mais atípica da variável em estudo.

A inclusão de diferentes dias no mesmo *cluster* de observações não implica necessariamente que a variabilidade entre eles não possa atingir valores com alguma dimensão. No entanto, as evoluções são muito mais diferenciadas como consequência de nos referirm

os a dias laborais ou de fim-de-semana ou por repercussão de determinados programas transmitidos nos canais concorrentes. Consequentemente, os valores de *Rat_Prog%* podem apresentar uma maior variabilidade como resultado das oscilações horárias dos programas e das durações programáticas e em menor grau de fortes alterações nos comportamentos de consumo televisivo para idênticos períodos horários no decorrer da semana. Contudo, apenas com a análise dos indicadores de duplicação das audiências poderemos avaliar o grau de estabilidade dos padrões de consumo televisivo.

Verificamos que, a nível médio, os dias laborais diferenciam-se em maior grau dos dias de fim-de-semana apenas nos períodos de maiores níveis de audiências.

Para o Canal TVI (Gráfico F.2.2.3.I.2), aplicando o algoritmo K Média, verificamos que no *cluster* 4 constam alguns dias de fim-de-semana ou semanais onde foram transmitidos programas desportivos essencialmente no canal RTP1. O *cluster* 1 inclui alguns dias essencialmente de fim-de-semana nos quais o canal RTP1 transmite programas de Fórmula 1 e o Festival da Eurovisão da Canção. Os dias correspondentes ao fim-de-semana distribuem-se entre o *cluster* 1 e 4 que representam evoluções mais atípicas e os dias semanais distribuem-se entre o segundo e terceiro *cluster* que apresentam evoluções médias muito próximas para a maior parte da emissão diária.

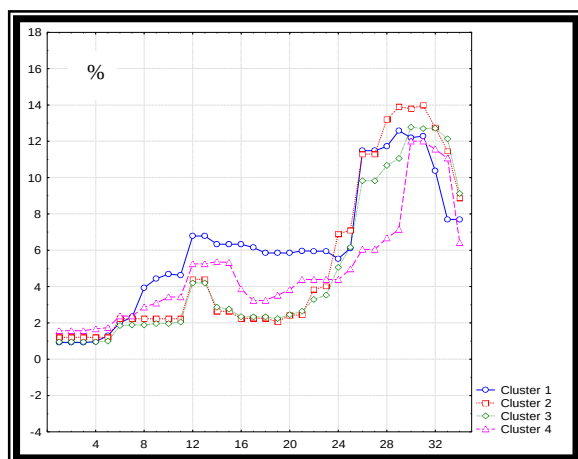


Gráfico F.2.2.3.I.2 – Clustering de séries temporais de valores de ratings para o canal TVI

Verificamos que existem evoluções diferenciadas entre o fim-de-semana e a semana, mas em cada um destes períodos menos equiparáveis e que contrasta com os resultados para o canal SIC.

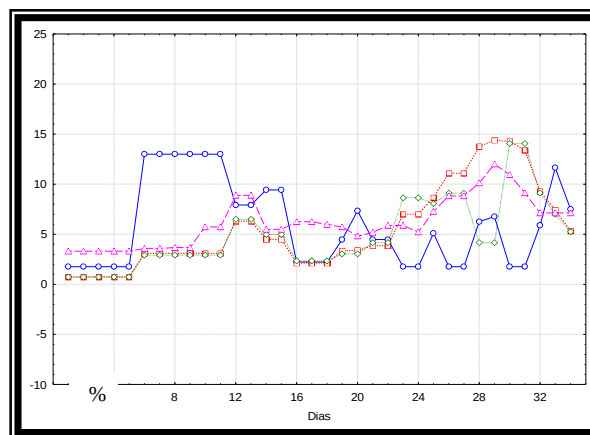


Gráfico F.2.2.3.I.3 – Clustering de séries temporais para o canal RTP1

Os anos 2000 a 2004, correspondem a um período de transição no qual a TVI consegue aumentar o seu *share* de audiência e começar a posicionar-se como um canal concorrente da SIC. Portanto, estes resultados encontram-se influenciados por também estar representado na amostra um período de consolidação da posição da TVI.

Para o Canal RTP1 (Gráfico.F.2.2.3.I.3) e para uma solução a quatro *clusters*, apenas o segundo se diferencia em maior grau dos restantes, incluindo dias onde foram transmitidos determinados jogos de futebol que envolvem os Clubes Sporting, Benfica e Porto. Verificamos uma maior similitude entre a evolução dos dados no decorrer da semana e ao fim-de-semana.

CANAL RTP1

Em suma, verificamos que o canal RTP1 apresenta uma evolução menos diferenciada entre os dias semanais e os de fim-de-semana, verificando-se evoluções mais atípicas de evolução dos *ratings* quando programas de grande popularidade são transmitidos nomeadamente os jogos de futebol.

CANAL SIC

O canal SIC apresenta uma nítida diferenciação entre o padrão de evolução semanal e de fim-de-semana, existindo alguma regularidade em cada um destes períodos. Alguns eventos desportivos dos canais concorrentes são os responsáveis por evoluções que se diferenciam significativamente da média, no entanto são em menor número face ao canal TVI.

CANAL TVI

A TVI é penalizada ou favorecida com a transmissão de eventos desportivos por parte dos canais concorrentes, existindo uma diferenciação consoante estes correspondem a jogos de futebol ou a outra tipologia de eventos, sejam eles desportivos ou não, como por exemplo, a Fórmula I ou a festivais da Eurovisão.

As evoluções semanais podem-se diferenciar, contudo verifica-se essencialmente ao nível do horário nobre e dos programas sucessores. O número de dias com evoluções mais atípicas é superior ao do canal SIC, uma vez que apresenta um posicionamento menos consolidado em parte do período considerado.

Existe uma notória divergência quanto à evolução e grau de variabilidade dos *ratings* programáticos e por faixas horárias para cada um dos canais considerados e quanto à respectiva propensão para se verificarem tendências que se afastem da média e que condicionam futuras análises de previsão. O mesmo programa apresenta repercussões diferenciadas em função do canal de transmissão, assim como a contra- programação dos canais concorrentes.

Verificamos que apesar das alterações que se possam verificar no decorrer das semanas quanto ao conteúdo programático existe alguma regularidade na ordem de grandeza dos *ratings* para determinados períodos horários e perante uma grelha de programação já estabelecida que não incluía determinados eventos, denotando em princípio uma certa fidelização da audiência ao meio televisivo em si, a programas e a determinados canais.

Os programas diferenciam-se quanto a serem mais variáveis em dias laborais ou ao fim-de-semana. Constatamos que os programas da noite apresentam, por vezes, uma maior estabilidade de *ratings* ao fim-de-semana. Com base nas análises das taxas médias de audiências para o total da TV, a maior estabilidade no decorrer dos anos associa-se aos períodos de lazer (pós laboral e ao fim-de-semana), logo a variabilidade ao nível dos *ratings* semanais deve-se, talvez, em maior grau a alterações nos horários de transmissão, na duração da exposição e ao grau de actividade na exposição televisiva e em menor grau aos factores de disponibilidade do painel.

Ignorando observações atípicas será essencialmente nos períodos de maiores audiências da noite que se podem verificar as maiores divergências no decorrer dos anos, assim como para os canais privados. Verificamos que para a totalidade dos canais e em especial para a RTP1 a transmissão de eventos desportivos pode provocar uma súbita muito significativa nos valores dos *ratings*. Para o canal SIC, os diferenciais de valores face à média observada são superiores ao fim-de-semana, mas em ambos os períodos não transcendem os 5%.

Para o canal RTP1 não existe uma grande diferenciação nestes períodos e os diferenciais de valores, também não transcendem os 5%, assim como para a TVI.

Verificamos que os factores temporais influenciam a evolução dos *ratings* e o seu grau de variabilidade, assim como o canal considerado. Consequentemente, no capítulo seguinte em vez de realizar o *clustering* dos dados que representam a totalidade da emissão diária, em primeiro lugar procede-se à sua segmentação, uma vez que no decorrer do dia a evolução dos *ratings* difere substancialmente.

II. Análises de Segmentação da variável *Rat_Período%*

Neste capítulo pretende-se verificar qual é a real subdivisão da emissão diária, atendendo à evolução dos dados. Com base nas análises exploratórias realizadas verificámos que existe alguma independência dos valores de *ratings* associados a períodos horários, face às alterações verificadas na grelha de programação. No entanto, o grau de regularidade encontra-se fortemente dependente do canal, dias considerados e períodos horários.

Neste capítulo procede-se à segmentação da emissão diária em função da variabilidade observada neste indicador, considerando a totalidade dos canais ou de forma isolada. Estas análises permitirão a identificação de horizontes temporais com distintos comportamentos de consumo televisivo para cada um dos canais e que também contribuem para a posterior previsão deste indicador. Verificamos que o desenvolvimento de modelos explicativos e de previsão das audiências implica, sem dúvida,

considerar a faixa horária a que um programa corresponde e, posteriormente em determinados contextos, também, a análise de diferentes cenários de contra-programação e de factores de planeamento. Um cenário de programação transmitido com uma maior frequência corresponderá em princípio a intervalos menos significativos de variabilidade de valores de *ratings*.

Consequentemente, pretende-se também avaliar em que medida os valores associados a períodos horários, minimizam ou não, a variabilidade associada a programas em concreto, sendo estes últimos valores penalizados como consequência das oscilações horárias programáticas. As taxas médias de audiências para o total da TV podem apresentar uma evolução diferente das taxas médias de audiências, no que diz respeito aos padrões de variabilidade e são função do canal considerado.

Segundo *Fernandes (2001)*, a emissão diária subdivide-se em quatro períodos horários: durante o dia, horário nobre, acesso ao horário nobre e fim da noite, designados de *carrefours* de programação. Pretendemos constatar se realmente se verifica esta segmentação da emissão diária atendendo à evolução e variabilidade dos índices de audiências. Construiu-se uma base de dados em que se associa a cada intervalo de meia hora para o total da emissão diária o valor de *Rat_Prog%(P)* que lhe está associado para cada um dos canais, em função dos programas que abrange. A base de dados simbólica correspondente traduz a variabilidade dos *ratings* para cada um dos períodos e por canal, podendo-se utilizar os dados referentes à totalidade da semana ou diferenciando o período semanal do fim-de-semana.

Do ponto de vista metodológico recorremos a algoritmos de *clustering* simbólicos, nomeadamente às Redes Neurais SOM e ao *Clustering* Divisivo.

Se o objectivo das análises se traduzir numa mera previsão de valores de *ratings* em determinadas faixas horárias, como foi demonstrado podemos recorrer apenas à análise das séries temporais ignorando os diversos factores subjacentes e respectivas interacções. Nesse sentido pretende-se avaliar quais são as abordagens metodológicas que produzem os melhores resultados.

Pretende-se a análise de *Rat_Período%*, sendo posteriormente necessário a sua afectação aos programas transmitidos.

Os resultados diferenciam-se em três itens: (1) análises para cada um dos canais; (2) análises considerando a globalidade dos dados, (3) análises para a globalidade dos dados, mas diferenciando o fim-de-semana e os dias laborais.

11a. Análises para cada um dos canais

Aplicando o algoritmo de *clustering* Simbólico divisivo aos dados do canal RTP1 para uma solução a quatro *clusters* obtemos uma variância explicada muito significativa no valor de 97% (Fig. F.2.2.3.II.1).

Para uma solução a quatro *clusters* obtemos conjuntos de faixas horárias que apresentam uma variabilidade de valores de *ratings* que se aproxima dos 5%. Se aumentarmos o número de *clusters* conseguimos reduzir esta amplitude de valores.

	Cluster 1		Cluster 4
	Cluster 2		Cluster 3

7		8		9		10		11		12

13	14	15	16				18

19	20	21	22	23	24

Fig. F.2.2.3.II.1 – Representação de segmentos horários para o canal RTP1

SEGMENTO_1

Para o período em análise as audiências são minimizadas em todo o período matinal até às 13H; no decorrer de toda a tarde, logo após a conclusão do programa sucessor ao espaço informativo (15H) até ao início do pré-prime-time (19H) e a partir das 23H (Cluster_1 com *Rat_Período_RTP1%* < 4,9%).

SEGMENTO_2

A faixa horária das 22H às 23H corresponde a concursos, como, por exemplo, ao programa “Um Contra Todos” incluindo-se no segundo *cluster* com os maiores valores de *ratings*. Apresenta valores de *Rat_Período_RTP1%* ε [9,62%; 14,75%] e associa-se ao período que decorre entre as 19H-19.30H uma vez que também corresponde a um período de transição na evolução das audiências.

SEGMENTO_3

O horário de maximização das audiências tem início às 19.30H prolongando-se até às 22H. O concurso “O Preço Certo” sem dúvida que pela sua popularidade funciona como ancora para o espaço informativo da noite. O Jornal da Noite e os seus programas sucessores incluem-se nos períodos que maximizam as audiências.

Faixas Horárias: 19.30H-22H

SEGMENTO_4

Este segmento corresponde aos períodos que apresentam o segundo intervalo de maiores valores de ratings e está associado ao espaço informativo da hora do almoço e ao início do programa que o sucede.

Recorrendo ao algoritmo de *clustering* divisivo simbólico podemos identificar faixas horárias comparáveis quanto à evolução deste indicador é proceder ao cálculo de um intervalo de variabilidade de *Rat_Periodo%_RTP1* para os períodos que cada *cluster* horário abrange.

Estes intervalos apresentam alguma amplitude de valores de *ratings*, mas que pode ser reduzido se aumentarmos o número de *clusters* considerados. Consequentemente, esta perspectiva de análise apresenta potencialidades. Verificamos que no período designado por final da noite, existem faixas horárias com diferenciadas evoluções e que por vezes até se aproximam de outros períodos horários que alguns autores apresentam como sendo divergentes.

Atendendo aos dados do canal TVI corresponde-lhe 92% de inércia explicada para uma solução a quatro *clusters* e de 95% para cinco *clusters*.

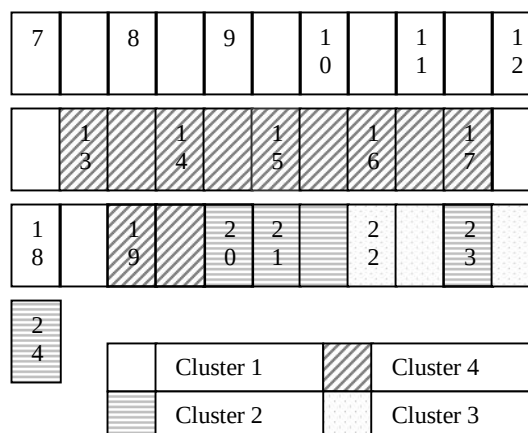


Fig. F.2.2.3.II.2 – Representação de *clusters* horários para o canal TVI

Para uma solução a quatro *clusters* chegamos a um diferencial de valores por volta de 3% de valores de *ratings* de *cluster* para *cluster* (Fig. F.2.2.3.II.2).

SEGMENTO_1

Os menores valores correspondem ao período matinal até às 13H e das 17.30 H-19H (*Rat_Periodo_TV1%*<4.37%) para o período em análise.

SEGMENTO_2 E SEGMENTO_3

Os *clusters* com maiores níveis de audiências estão associado ao período que decorre entre as 20H e as 24H, mas com duas diferentes tipologias de evolução. O Segmento_2 abrange os períodos 20-22H e 24-24.30H.

Para uma solução a cinco *clusters*, o primeiro e o quarto agrupamentos mantêm a sua composição e as restantes observações subdividem-se da seguinte forma: (1) 20H-22H (*Cluster 5* - *Rat_Período_TVI%* $\in$ [8,6; 9,85]); (2) 22H-23H e 23.30H-24H (*Cluster 3* - *Rat_Período_TVI%* $>$ 9,85%) (3) 23H-23.30H e 24H-24.30H (*Cluster 2* - *Rat_Período_TVI%* $\in$ [6,95 - 8,6%]).

As audiências maximizam-se para o terceiro *cluster*, seguindo-lhe o quinto e o segundo. No decorrer da noite existe uma maior oscilação nas audiências, maximizando-se no período compreendido entre as 22H e as 23H (primeira telenovela da noite) e no período das 23H-23.30H (segunda telenovela da noite). O período matinal e a tarde apresentam os menores valores. Com o aumento do número de *clusters* conseguimos reduzir o intervalo de variabilidade de valores de *ratings*.

Atendendo aos dados do canal SIC para uma solução a cinco *clusters* a percentagem de variância explicada é de 96,6% e de 90% para quatro *clusters*. Os maiores valores de *ratings* estão associados às faixas horárias 22H-23H e aos dois espaços informativos do dia.

O segundo *cluster* com maiores valores corresponde desde as 16.30H até às 20H (incide em telenovelas durante os dias laborais e em talkshows); entre as 21H e as 22H (programas sucessores ao espaço informativo e que podem anteceder a principal telenovela da noite) e após as 23H (sucessivas telenovelas da noite). O período da manhã até às 12H corresponde ao que apresenta menores valores de *ratings*, seguindo-se as faixas horárias 12-13H e 14H-16.30H.

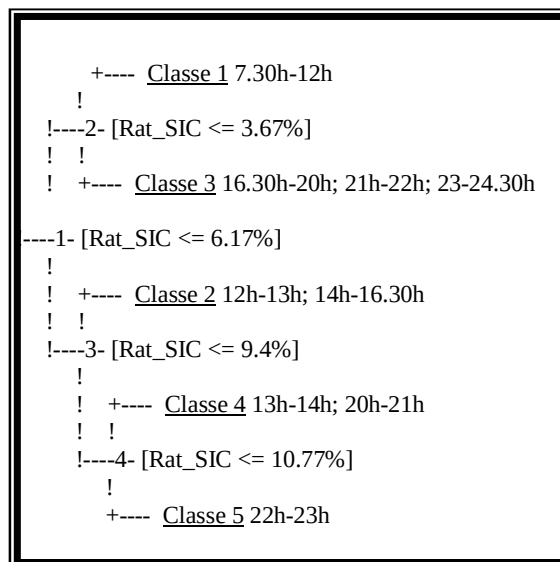


Fig. F.2.2.3.II.3 – Representação dos *clusters* horários para o canal SIC

SIC	TVI
22h-23h ($>10,77\%$)	22h-23h + 23.30-24h ($>9,85\%$)
13h-14h; 20h-21h ($9,4\% - 10,77\%$)	20h-22h
16.30h-20h; 21h-22h; 23h_FIM ($3,67\% - 6,17\%$)	23h-23.30h + 24h- 24.30h
12h-13h; 14h-16h ($<9,4\%$)	13h-17.30h, 19h-20h ($<6,95\%$)
7.30h -12h ($<3,67\%$)	7.30h-13h; 17.30h- 19h ($<4,37\%$)
RTP1	
19.30h-22h ($>14,7\%$)	
19h-19.30h, 22h-23h ($<14,7\%$)	
13h-15h e 23h-23.30h ($<9,6\%$)	
7h-13h; 15h-19h; 23.30h- FIM ($<4,9\%$)	

Tabela F.2.2.3.II.1 – Representação de faixas horárias por ordem crescente de valores de *Rat_Período%*

Na *Tabela F.2.2.3.II.1* ordenam-se os períodos horários por ordem crescente de *Rat_Período%*. Para o canal estatal o período com maiores índices de audiência começa mais cedo e abrange um maior horizonte temporal, incluindo o espaço informativo da noite. Para os canais privados os maiores índices estão associados às telenovelas da noite com início por volta das 22H e os segundos maiores valores correspondem a ambos os espaços informativos ou apenas ao programa da noite. A RTP1 também apresenta maiores valores de índices nos períodos adjacentes ao seu horário nobre. O canal estatal diferencia-se dos canais privados por apresentar períodos mais alargados com uma evolução indiferenciada ao nível dos *ratings*. Foram realizadas idênticas análises para os anos remanescentes (2004-2008) e a aplicação de algoritmos simbólicos a dados de períodos horários permite a identificação de períodos horários com uma variabilidade mais restrita. Por aplicação do algoritmo DIV será possível calcular um intervalo de confiança para determinados *clusters* de valores de *Rat_Período%*. Estes resultados são válidos para uma determinada grelha de programas, devendo ser recalculados para outros horizontes temporais que se afastem do consumo televisivo médio ou quando se verifiquem consideráveis alterações na estruturação da grelha de programação.

IIb. Resultados das Análises considerando a totalidade da amostra

O algoritmo de *clustering* simbólico divisivo foi anteriormente aplicado a cada um dos canais, permitindo a identificação dos períodos com maiores e menores índices de audiências atendendo à variabilidade observada neste indicador.

Considerando simultaneamente a totalidade dos canais e por aplicação dos algoritmos de *Clustering* simbólicos de Redes Neurais SOM e de *Clustering* Divisivo pretende-se a segmentação da emissão diária atendendo à evolução global dos *ratings*. Iniciaram-se as análises não se diferenciando os períodos horários em função do fim-de-semana e dos dias laborais.

□ Resultados que decorrem da aplicação das Redes Neurais SOM Simbólicas

Segundo as Redes SOM simbólicas, o melhor resultado corresponde a quatro *clusters*, respeitando a condição de que n que representa o número de períodos-34 a dividir por m (n° de *clusters*) deverá ser superior ou igual a cinco. Se incrementarmos o número de *clusters* obtemos conjuntos sem observações. Utilizou-se como metodologia o *StochApprox* que corresponde a uma generalização do método clássico das SOM e a medida L2 e Hausdorff.-Type para o cálculo das distâncias e para o cálculos dos protótipos os vértices médios. Para $x_{uj}=[\alpha_{uj},\beta_{uj}]$ e $x_{vj}=[\alpha_{vj},\beta_{vj}]$ a medida L2 é dada por

$$\delta_j(x_{uj}, x_{vj}) = (\alpha_{uj} - \alpha_{vj})^2 + (\beta_{uj} - \beta_{vj})^2.$$

O *Segmento 1X1* corresponde aos programas que antecedem o espaço informativo da tarde e ao início da tarde, ou seja, às faixas horárias 11.30H-13H e 15H-16.30H. O *Segmento 1X2* diz respeito ao início da manhã (7H-11.30H) e o *Segmento 2X2* corresponde aos períodos sucessores do espaço informativo da tarde, ao fim da tarde até ao início do pré-prime-time e ao final da noite (14H-15H, 16.30H-19H, 23.30H-24.30H). Por último o *Segmento 2X1* corresponde ao espaço informativo da tarde (13H-14H) e às faixas horárias que decorrem entre as 19H e as 23.30H.

Segmento 1X1

Ao longo de toda a semana e a partir das 11.30H verifica-se uma acentuada subida em *Tvr_Período%*, sendo esta subida mais notória para os *ratings* de fim-de-semana e que culmina com o espaço informativo da tarde. Estes resultados são consistentes com as análises gráficas realizadas à variável *Tvr_Período%* e a *Rat_Prog%*.

Esta faixa horária é agregada ao período das 15H-16.30H, uma vez que também corresponde a um período de transição de valores de *ratings* em sentido inverso, mas que ao nível de valores de *Tvr_Período%* corresponde a um patamar mínimo de valores após o qual tem início uma nova tendência ascendente.

Segmento 2X2

Agrega os períodos de transição adjacentes ao que tradicionalmente se define como prime time e da hora de almoço e que apresentam evoluções semelhantes ao nível de *Tvr_Período%*.

Segmento 1X2

Abrange os períodos que minimizam os valores de *ratings*, ou seja, o período matinal até às 11.30H.

Segmento 2X1

Inclui o espaço informativo da tarde e o horário nobre entre as 19H e as 23.30H. O pré-prime-time equipara-se ao nível de *Tvr_Período%* com a hora do almoço, subindo posteriormente até atingir o seu nível máximo entre as 20H-22.30H e posteriormente voltando aos níveis iniciais.

Uma vez identificados os segmentos horários analisam-se os valores dos correspondentes protótipos e os valores de *Rat_Período%*.

Por análise do gráfico que representa a variabilidade dos protótipos dos *clusters* associados aos dois canais privados verificamos que o segmento 2X1 que inclui os períodos de maiores audiências, apresenta, também, a maior amplitude de valores de *ratings* e o matinal obviamente os menores. Os protótipos foram calculados como a média dos vértices de cada elemento que o constituem. A amplitude de valores para todos os períodos horários é menor para o canal SIC em relação aos valores da TVI. Seria desejável uma maior separação entre os valores dos protótipos, mas isto denota a evolução continua dos *ratings*, sendo as transições entre períodos horários muito subtis o que torna difícil a sua segmentação. Na *Tabela F.1.2.2.3.III.2* constam os valores dos *ratings* de programas, e não de períodos horários, para cada um dos *clusters* horários e canais e a correspondente variabilidade de valores.

RATINGS DE PROGRAMAS			
	Rat_SIC%	Var_Rat_ SIC	Rat_RTP%
1x1	0,50-4.10	3,7% (3,4%)	0,4-6.30
1x2	1.70-7.80	6% (5,9%)	1.60-6.30
2x1	1.90-17.20	15% (14,3%)	0.20-30.70
2x2	1.80-14.20	12% (9,5%)	1.10-10.3
	Var. Rat_RTP%	Rat_TVI %	Var. Rat_TVI%
1x1	6% (5,9%)	0.60-5.20	4,6% (4,5%)
1x2	4.7% (4,7%)	1.20-7.7	6,5% (6,5%)
2x1	30% (21,3%) (12,3%)	1.90-17	15% (14,8%)
2x2	9,2% (7,1%)	1.30-2.30	13% (13%)

Tabela F.1.2.2.3.III.2 – Valores de *ratings* para cada segmento horário identificado com as Redes Neurais SOM simbólicas

Em parênteses constam os valores, excluindo as observações que correspondem aos programas desportivos. As variabilidades máximas estão associadas aos *clusters* 2X1 e 2X2, diferenciando-se de forma máxima para o *cluster* 2X1 do canal estatal. A variabilidade de valores para os canais privados é muito semelhante. Excluindo programas desportivos os valores de variabilidade reduzem-se para os valores entre parêntesis, sendo essencialmente mais significativas para o canal estatal. A descida é sobretudo notória para a RTP1 com o valor de 9% e maior ainda se excluirmos também os programas sucessores dos eventos desportivos que beneficiam dos efeitos de herança gerados nestes contextos, com o valor de 17,7%, ficando assim ao mesmo nível de variabilidade dos canais privados. Apesar de determinados jogos de futebol reduzirem para metade os índices de audiências dos canais concorrentes, excluindo estes programas, a variabilidade global reduz-se apenas ligeiramente para os canais privados, ficando o seu efeito atenuado na globalidade dos dados.

□ Resultados segundo o algoritmo imbólico de *Clustering Divisivo*

Se utilizarmos o algoritmo de *clustering* divisivo simbólico aplicado aos dados em que os intervalos de *ratings* para cada um dos canais estão associados à totalidade das observações, os *clusters* resultantes diferem dos obtidos anteriormente. A Fig. F.2.2.3.II.4 traduz as variáveis e os valores utilizados nos *splits*. Para uma solução a dois *clusters* a percentagem de inércia explicada apresenta o valor de 70%, com a separação do período que inclui as horas que decorrem entre as 19H e as 23H das restantes observações.

Cluster 1 7HM 8H 8HM 9H 9HM 10H 10HM 11H 11HM 12H 12HM 14H 14HM 15H 15HM 16H
Cluster 2 19H 19HM 22H 22HM
Cluster 3 13H 13HM 16HM 17H 17HM 18H 18HM 23H 23HM 24H
Cluster 4 20H 20HM 21H 21HM

+----	Cluster 1
!	
!----	2- [Rat_SIC% <= 6.17%]
!	!
!	+----
!	Cluster 3
!	
!----	1- [Rat_Rtp% <= 9.62%]
!	
!	+----
!	Cluster 2
!	!
!----	3- [Rat_Rtp% <= 17.00%]
!	
+----	Cluster 4

Fig. F.2.2.3.II.4 – *Clustering* divisivo dos períodos horários e para uma solução a 4 *clusters*

O canal estatal corresponde ao que apresenta maiores oscilações e transições mais bruscas na evolução dos *ratings*, por conseguinte foi a variável utilizada no primeiro *split*. O *cluster 4* corresponde aos períodos que maximizam as audiências. Para uma solução a três *clusters* obtemos 80% de inércia explicada com a seguinte constituição de *clusters* (7H-13H; 14H-16.30H, *Rat_RTP%*<9.6% e *Rat_SIC%*<6.17%) e (13H-14H; 17H-19H; 23H-FIM, *Rat_RTP%*<9.6% e *Rat_SIC%*>6.17%). Para uma solução a quatro *clusters* a inércia explicada é de 88% e o *cluster* com as maiores audiências para a solução anterior é desagregado em 19H-20H e 22H-23H, *cluster 2*, 20H-22H, *cluster 4*.

O primeiro segmento corresponde aos períodos que tradicionalmente são classificados como pré-prime-time e pós-prime-time, onde assistimos ao início de uma transição na dimensão dos *ratings* e em *Tvr_Período%*. O segundo corresponde ao período designado de horário nobre, onde as audiências e *Tvr_Período%* atingem o seu máximo. Neste caso os *ratings* da RTP1 são a variável mais utilizada, correspondendo o primeiro *split* ao valor 9.62% e o segundo a 17%. Para os menores valores das audiências estatais este algoritmo utiliza os *ratings* da SIC com o valor de 6% a definir a fronteira entre as classes. Neste caso o pré-prime-time e o início da noite (*cluster 2*) é separado do prime-time (*cluster 4*). O *cluster* três corresponde à hora do almoço, ao período que decorre entre as 16.30H e as 19H e após as 23H. O *cluster 1* corresponde a todo o período da manh, e após o espaço informativo da tarde até às 16.30H, ou seja, aos níveis mais reduzidos das audiências. Para uma solução de cinco *clusters* o acréscimo na variabilidade explicada pode ser considerado marginal.

A solução proporcionada pelas Redes Neuronais SOM parece ser mais apropriada e lógica face à que resulta da aplicação do algoritmo DIV.

III. Resultados das Análises Diferenciando o Fim-de-semana e os Dias Laborais

Diferenciando as análises em função do fim-de-semana e dos dias laborais e por aplicação das Redes SOM simbólicas obtemos uma solução a seis *clusters* e aplicando como medida de cálculo das distâncias a L1. Para $x_{uj}=[\alpha_{uj},\beta_{uj}]$ e $x_{vj}=[\alpha_{vj},\beta_{vj}]$ a medida L1 é dada por $\delta_j(x_{uj},x_{vj})=|\alpha_{uj}-\alpha_{vj}|+|\beta_{uj}-\beta_{vj}|$.

Com base na análise dos protótipos dos *clusters* podemos agregar os segmentos 2X2 e 3X2, uma vez que lhes corresponde uma acentuada descida nos *ratings* para o canal estatal e maiores valores para os canais privados com variabilidades equiparáveis. O final da noite e o período pós espaço informativo da hora do almoço apresenta valores superiores para os canais privados em ambos os períodos e a tarde apenas ao fim-de-semana. Os segmentos 2X1 e 3X1 também se podem agregar observando-se variabilidades semelhantes para os vários canais e ligeiramente superiores para o segmento 2X1.

O segmento 1X1 apresenta valores muito superiores para o canal estatal face aos restantes, correspondendo à hora do almoço e ao pré-prime-time. No segmento 1X2, que corresponde aos períodos de maiores audiências, verificamos que os *ratings* maximizam-se para todos os canais, sendo o canal estatal o mais penalizado, uma vez que a tendência ascendente tem origem em períodos mais prematuros.

Segundo as últimas análises SOM, a emissão diária é dividida nos seguintes intervalos: o fim-de-semana até às 13H (Período 1); 13H-14H com 19H-23H (Período 2); entre as 14H e as 19H e após as 23H (Período 3). A emissão diária no decorrer da semana é subdividida no período matinal até às 13H (Período 1); das 13H às 15H (Período 4); das 15H-19.30H e a partir das 23.30H (Período 3) e 19.30H- 23.30H (Período 2).

Considerando os valores de programas em concreto o canal estatal é o que apresenta uma maior variabilidade ao nível dos *ratings* a partir da hora de almoço.

Os canais privados são mais estáveis e obviamente incrementam a variabilidade observada nos períodos de maiores audiências. Quanto aos valores de *shares*, as margens são muito elevadas para todos os canais e períodos considerados. Os valores de variabilidade de *ratings* são menores em relação aos que resultam da aplicação das redes SOM aos dados que englobam a semana e o fim-de-semana, mas ainda apresentam valores consideráveis para os segmentos 2X1 e 2X2. Desagregando o *cluster* 2X1 em 13H-14H e em 19H-23.30H verificamos que o primeiro período tem valores menos variáveis em relação ao segundo.

Com base nas análises de segmentação realizadas, as Redes SOM produziram os melhores resultados.

Nas análises posteriores adopta-se a divisão que resulta da análise da totalidade dos dados ou que diferencia os dados semanais e os dados de fim-de-semana. As previsões de *ratings* são função das faixas horárias, mas intervalos horários de meia hora são insuficientes para uma posterior avaliação da repercussão de factores de planeamento. Consequentemente, os segmentos horários que resultam da análise da variabilidade de *ratings* são os mais adequados para o desenvolvimento das análises propostas.

IV. Considerações sobre as análises de previsão de valores de *Ratings*

No capítulo anterior tentou-se prever as taxas médias de audiências para o total da TV recorrendo a metodologias não paramétricas, no entanto os algoritmos de *clustering* acabaram por produzir melhores resultados.

Neste capítulo também se pretende recorrer a metodologias de Regressão não paramétrica para tentar prever os valores de *ratings* recorrendo apenas às suas séries temporais univariadas e aos valores parciais.

Simultaneamente identificam-se quais são os factores responsáveis pela maior variabilidade dos valores de *ratings*. Constatámos que nem todos os programas que apresentam maiores oscilações quanto aos períodos horários em que são transmitidos, também apresentam grandes margens de variabilidade de valores de *ratings*.

Para o canal TVI os programas com maiores margens de valores de *ratings* correspondem aos espaços informativos da noite e às duas telenovelas e séries da noite. Correspondem essencialmente a programas de Segunda a Sexta, o que podemos atribuir a uma maior instabilidade nos padrões de disponibilidade do público face às condicionantes laborais e sem dúvida à transmissão de eventos desportivos nestes períodos.

Para o espaço informativo da noite e semanal se excluirmos os programas que tiveram como contra-programação os eventos desportivos os valores de variabilidade reduzem-se para 4,2% (12.9%-8.4%) e ao fim-de-semana para 4,6% (7.8%-12.4%). Ainda que se alterem os horários dos programas dos restantes períodos horários ou de fim-de-semana, estas alterações não se repercutem de forma tão significativa, uma vez mais como consequência da audiência estar mais regularmente disponível.

Os programas do canal SIC que apresentam uma maior variabilidade de valores, também não apresentam necessariamente maiores alterações nos períodos horários em que são transmitidos. A maior variabilidade corresponde ao espaço informativo da noite da semana e ao espaço informativo da tarde de fim-de-semana; ao programa de humor da noite ao longo de toda a semana e a filmes da noite ao fim-de-semana. O programa BBC Vida Selvagem, ou outros de idênticas características apresentam *ratings* mais estáveis no intervalo (4.20%-7.30%), apesar da sua grande oscilação horária. Para os restantes programas os seus intervalos de *ratings* não superam os 4% de variabilidade.

Para o espaço informativo da noite semana podemos atribuir a sua variabilidade (6,6%) aos eventos desportivos do canal estatal e que o penalizam (uma vez excluídos reduz-se este valor para 3,4%). Ao fim-de-semana excluindo esses programas a variabilidade é de 3.8%. (valores de *ratings* no intervalo 8.8%-12.6%).

Para o espaço informativo da tarde ao fim-de-semana sucede o inverso. A transmissão em simultâneo de dois programas de Fórmula1 na RTP favorece este canal, gerando observações atípicas de 13,7% e 11,2%. Excluindo estes programas a variabilidade é de apenas 1,1%.

Se considerássemos apenas os programas do período de maiores audiências a variabilidade seria de apenas 2,7%.

O programa de humor do horário nobre (S/FS) com variabilidade muito acentuada reduz o seu valor para 5,2% durante a semana e para 2,3% ao fim-de-semana excluindo os eventos desportivos. A variabilidade associada à primeira telenovela da noite não é possível reduzir.

Para o canal estatal RTP1 a maioria dos programas apresentam valores de variabilidade aproximadamente de 2%. Os valores são maximizados pelo Preço Certo (FS) e seguem-se os espaços informativos. Para o espaço informativo da noite da semana existem duas observações que minimizam atipicamente os seus valores e que correspondem a dois dias em que o programa foi transmitido mais cedo uma vez que lhe sucedia o festival da Eurovisão da Canção e um jogo de futebol. Existe também a situação inversa em que o programa é transmitido num horário mais tardio por lhe anteceder um jogo de futebol o que favoreceu significativamente os valores dos seus *ratings*.

Excluindo estes programas a variabilidade reduz-se para 2,9% (valores no intervalo 10%-12,9%). Para o espaço informativo da noite de fim-de-semana a variabilidade é de 2,1% se considerarmos apenas os programas que também têm como contra-programação os espaços informativos dos restantes canais (10%-12,1%), subindo para 5,8% por ser transmitido em outros horários. Os menores valores estão associados às transmissões que antecedem as 20H. No que diz respeito ao espaço informativo de fim-de-semana reduzimos para 3% a variabilidade dos seus valores se excluirmos os programas desportivos.

Portanto, uma vez mais as alterações nos horários de transmissão e principalmente a programação adjacente influencia significativamente os resultados. Para o canal RTP1 os programas beneficiam de ser antecidos por programas de maior popularidade ainda que isso implique adiar o seu início, mas são penalizados quando são sucedidos por programas de grande projecção se estes últimos o anteciparem.

Verificamos que no decorrer das semanas em dias laborais existe uma forte semelhança na evolução dos *ratings* até ao pré-prime-time, consequentemente por via do *clustering* das séries temporais talvez seja viável a previsão dos seus valores para os períodos que apresentam uma maior regularidade ao nível dos *ratings*. Nos períodos de maiores audiências verifica-se a maior discrepância e será necessário a análise dos efeitos de herança e dos fluxos de audiências entre os canais, assim como dos restantes factores estruturais que condicionam as audiências televisivas.

A lealdade programática, tipológica e ao canal, as variáveis temporais, os padrões de disponibilidade e a contra-programação são algumas das variáveis que influenciam de forma variável no decorrer do tempo os valores dos *ratings* programáticos.

A evolução dos *ratings* a curto prazo pode ser consideravelmente semelhante se excluirmos programas de maior popularidade.

O prime-time e a noite correspondem aos períodos onde os valores dos *ratings* podem apresentar uma maior variabilidade de valores, especialmente se forem transmitidos programas mais populares.

Para a previsão de valores de *ratings* ajustaram-se polinómios do quarto e quinto grau que apresentam consideráveis valores dos coeficientes de determinação com um valor máximo de 93% para um polinómio de grau 5 e com o valor máximo de 82% para um polinómio de grau 4 na amostra considerada. Alternativamente, recorreu-se às Médias Móveis cujos desvios aos valores observados também é pouco significativo.

Consequentemente, para realizar a previsão de valores de *ratings* as Regressões não lineares apresentam potencial de aplicação, nomeadamente as polinomiais, as regressões trigonométricas ou as regressões locais polinomiais. Também, podemos realizar o *clustering* das séries e considerar como previsão os protótipos representativos de cada um dos *clusters*. A área do *Data Mining* Temporal abrange inúmeros algoritmos que promovem o cálculo de dissimilaridade, classificação e *clustering* de séries temporais.

F.3. ANÁLISE DE ÍNDICADORES DE DUPLICAÇÃO DAS AUDIÊNCIAS: A LEALDADE PROGRAMÁTICA

This study aims at finding the extent of audience overlap between subsequent programs and comparing them to the results of American and British studies. ... It is rather difficult to get a consistent conception of detailed repeat viewing patterns, largely due to variations in study designs and continuing changes in the media environment. ... The response rates in diary samples are somewhat lower than in people-meter samples. The lower willingness to participate in a diary sample is presumably due to higher demands in terms of time spent in filling out diary sheets. ... A contrasting study in less industrialized countries may be helpful.

Zubayr (1999)

People-meter panels are for the first time providing, continuous week-by-week information on individual's television viewing in the United States. This provides the opportunity for more detailed analysis of viewing patterns, ... since people-meter panels are here to stay, there will be plenty of further data on repeat-viewing and other aspects of audience flow week-by-week for individual viewers. ... Broadly similar figures are thrown up by very different measurement techniques.

Ehrenberg & Wakshlag (1987)

The people-meter technology appears to be the television audience measurement tool of the future. ... The people meter is expected to replace an existing television audience measurement system that has served the television industry well over three decades. ... For the standard four-week survey period, the people-meter sample is considerably more reliable than a diary sample with the same total number of households.

Soong (1988)

Webster and Litchy (1991) observed that studies utilizing data from people-meter would show lower levels of repeat viewing than earlier diary-based research, since "diaries are often filled out after the fact, and they may reflect a recollection of consistent viewing, even if actual use was more sporadic. ... they argued that structural factors provided the most comprehensible means of developing meaningful research questions and producing practical results. ... more sophisticated techniques may be required to isolate the separate effects of this variables (scheduling) on repeat viewing.

Sherman (1995)

F.3.1. Introdução

Um dos factores que influencia a evolução dos *ratings* programáticos corresponde ao comportamento de lealdade programática o qual, segundo a literatura internacional da especialidade anglo-saxónica, também se pode denominar de "repeat viewing".

Os trabalhos realizados têm procurado caracterizar a tendência que alguns telespectadores apresentam no sentido de seleccionar episódios consecutivos de uma série programática, avaliando simultaneamente a influência de alguns factores estruturais a este nível. Os autores recorrem essencialmente a programas que apresentam uma calendarização regular e que são transmitidos em horários fixos (Zubayr, 1999).

Segundo Wakshlag (1985), seria de esperar elevados níveis de duplicação das audiências, uma vez que as eventuais preferências dos telespectadores por determinados programas e canais incentiva-os a assistir aos episódios subsequentes e estes últimos, por vezes, também apresentam uma reduzida variabilidade dos seus valores de *ratings* (Barwise, 1986; Ehrenberg & Wakshlag, 1987).

Os trabalhos iniciais foram essencialmente realizados nos USA e em Inglaterra, onde se constatou que a maioria das tipologias programáticas apresenta valores de duplicação das audiências que se aproximam dos 50% (Barwise, Ehrenberg & Goodhardt, 1982). Algumas análises posteriores demonstraram que os níveis de duplicação são na realidade muito menores. Por exemplo, em Inglaterra onde o número de canais na altura era reduzido, apenas metade da audiência assiste a episódios consecutivos de determinado programa (Goodhart, Ehrenberg & Collins, 1987). Também, na Alemanha (Zubayr, 1999), apenas um terço do painel ou uma percentagem ainda menor assiste a episódios consecutivos de um programa. De acordo com Ehrenberg e Wakshlag (1987), estes índices apresentavam na altura um valor médio no valor de 24% considerando uma amostra de dados dos USA. Na opinião de Zubayr (1999), os contextos e as metodologias influenciam os resultados, e, logo, dificilmente as conclusões podem ser generalizadas: "It is rather difficult to get a consistent conception of detailed repeat viewing patterns, largely due to variations in study designs and continuing changes in the media environment."

A estabilidade que os valores de *ratings* programáticos ou de valores de *ratings* referentes a períodos horários exibem, por vezes, a nível nacional em determinados horizontes temporais, também, *a priori* nos poderia levar a depreender que a duplicação de audiências deverá apresentar valores consideráveis em determinados contextos.

Todavia, os *ratings* ponderam os telespectadores relativamente ao tempo de contacto com os suportes, consequentemente ainda que a duplicação de audiências fosse de 100%, os valores de *ratings* não seriam necessariamente equivalentes, assim como menores valores de duplicação podem produzir uma maior semelhança dos valores de *ratings* de suportes adjacentes ou próximos, como consequência das variações observadas na duração da exposição dos telespectadores. Portanto, as duas variáveis não se encontram necessariamente correlacionadas, isto é, a partir da evolução de uma variável pouco podemos concluir a respeito da evolução da segunda variável.

Outra das limitações que grande parte dos trabalhos internacionais apresenta consiste em não conseguir enquadrar estes comportamentos na dicotomia de exposição televisiva activa vs. passiva, uma vez que os programas em análise são essencialmente transmitidos em faixas horárias equivalentes (Zubayr, 1999). Com base nas análises realizadas aos valores de *ratings* verificámos que um bom desempenho de um programa pode reduzir em parte a variabilidade observada nos valores dos seus *ratings* decorrente das frequentes alterações horárias a que estão sujeitos e em especial ao fim-de-semana. Em períodos de lazer mais facilmente o público-alvo de determinado programa poderá aguardar pela sua transmissão em horários mais tardios. Consequentemente, leva-nos a pressupor que poderá ser significativo o valor da duplicação de audiências em programas que apresentam um maior desempenho, existindo em especial nos períodos de lazer alguma actividade subjacente para parte do painel de audimetria.

No entanto, as frequentes alterações horárias, em geral, originam significativos intervalos de variabilidade para os principais indicadores de audimetria. A duplicação de audiências no período semanal deverá resultar em maior grau de uma exposição ritualística do painel, isto é, de uma exposição onde os telespectadores recorrem ao *media* para preencher os seus tempos livres ou vazios e que se encontra fortemente condicionada pela sua disponibilidade, entre outros factores estruturais. Na literatura da especialidade, como factores responsáveis dos menores valores de duplicação observados é sugerida a disponibilidade (Barwise, Ehrenberg & Goodhardt, 1982; Ehrenberg & Wakshlag, 1987) e as características de planeamento (Goodhart, Ehrenberg & Collins, 1987; Kamerer, 1987; Webster & Wang, 1992).

Ehrenberg e Wakshlag (1987) justificam os reduzidos níveis de duplicação com a irregularidade do consumo televisivo. Por exemplo, é salientado que, pelo menos 40% dos telespectadores não vê televisão quando se transmite um episódio de um programa semanal. Logo, apesar da regularidade, por vezes, verificada na evolução dos índices de audiência, uma grande percentagem da audiência não será certamente a mesma. Outros trabalhos (*Barwise, 1986; Barwise, Ehrenberg & Goodhardt, 1982; Ehrenberg & Wakshalag, 1987; Webster & Wang, 1992*) têm procurado relacionar os índices de duplicação com os valores dos índices de audiência dos programas envolvidos e, em geral, defendem que quanto maior for a popularidade de um programa maior será, também, a tendência para seleccioná-lo com uma maior regularidade.

No entanto, nem sempre os *ratings* representam correctamente o grau de envolvimento do público com a grelha de programação. Segundo *Webster e Wang (1992)*, os índices de duplicação são superiores no decorrer dia, superando os valores observados no período de prime time, em parte devido aos factores de planeamento, tais como: (1) Durante o dia existe maior número de programas com o formato de séries, por exemplo, as telenovelas; (2) Durante o dia existe uma maior regularidade no padrão do consumo semanal; (3) Durante o dia existe uma audiência mais continuamente disponível. Por outro lado, certas tipologias programáticas contribuem para uma maior duplicação das audiências, como, por exemplo, a estrutura narrativa das telenovelas convertem os seus programas em uma das tipologias que apresentam os maiores valores. Entre as restantes tipologias programáticas, não têm sido identificadas grandes diferenças (*Barwise, 1986; Barwise, Ehrenberg & Goodhart, 1982; Headen, Klompmaker & Rust, 1979; Webster & Wang, 1992*).

A nível nacional as frequentes transmissões de telenovelas nos períodos de maiores audiências, poderá atenuar o diferencial de valores observados nestes períodos face aos valores observados no decorrer das restantes faixas horárias. Nas análises realizadas às taxas médias de audiências para o total da TV, constatámos que este indicador apresenta uma maior estabilidade nos períodos de lazer, ou seja, ao fim-de-semana e em horários pós laboral. Se considerarmos a sua variabilidade bimensal, a maior disponibilidade do público ao fim-de-semana gera uma maior variabilidade ao longo de todo o dia até ao período de *lead in* ao

espaço informativo da noite, após o qual existe uma audiência total mais estável. Ao fim-de-semana, a existência de uma audiência sem condicionantes laborais reduz a sua variabilidade ao nível dos períodos de maiores audiências da noite. Uma vez que no decorrer da semana quem exerce uma actividade profissional não poderá ver televisão até às 18H-19H, a evolução de *Tvr_Período%* apresenta uma maior estabilidade neste período, ou seja, existe uma audiência mais regularmente disponível.

Consequentemente, a nível nacional, os índices de lealdade programática nos períodos de maiores audiências não têm que ser necessariamente muito inferiores aos observados no decorrer do dia. Além do mais, quanto mais elevado for o número de telespectadores expostos a determinado programa maior será a probabilidade de se verificarem maiores valores dos índices de duplicação de audiências. A nível nacional, a frequente transmissão de telenovelas nacionais e brasileiras nos períodos de maiores audiências da noite, também, poderá contribuir para a sua maximização, tornando-os equiparáveis relativamente aos valores observados no decorrer da tarde e da manhã, ainda que os *ratings* sejam mais estáveis nestes períodos. Verificámos que a maior variabilidade dos valores de *ratings* no período da noite, resulta em maior grau das frequentes alterações horárias, durações programáticas e da contra-programação que gera maiores níveis de exposição ao painel e em menor grau das alterações associadas às taxas médias de audiências para o total da TV dos suportes publicitários.

A maioria dos trabalhos internacionais identificados não são recentes, com excepção do trabalho realizado por *Xiaoling e Hing-Po Lo (2007)* que recorre a dados de audimetria chineses. Os trabalhos existentes incidem maioritariamente em bases de dados dos USA e Britânicos e em *Zubayr (1999)* consta uma análise comparativa face aos resultados anglo-saxónicos recorrendo a dados Alemães. Os trabalhos, também, incidem nos períodos de maiores audiências, por aplicação de metodologias estatísticas clássicas e utilizando como variável dependente no máximo as audiências conjuntas de cada par programas, a maioria dos quais consecutivos e considerando apenas os dados para o universo televisivo.

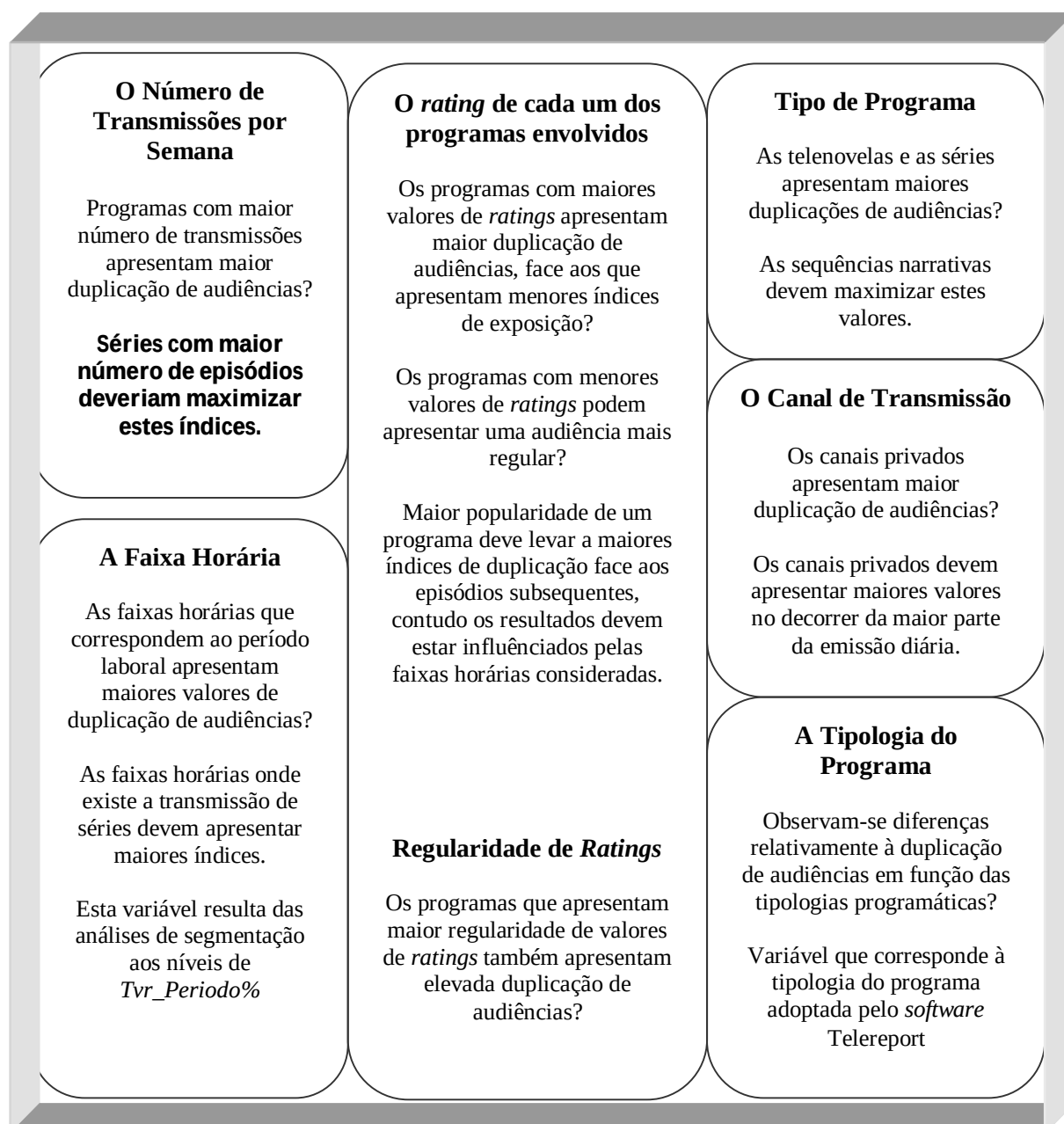
Na actualidade, a existência de um crescente número de alternativas lúdicas na área dos *media*, e não só, e a própria estruturação da grelha de programas que levou a uma maior fragmentação das audiências poderá reduzir a dimensão dos índices de duplicação. Os menores valores que se esperam, na actualidade, a nível internacional indiciam uma audiência talvez mais activa, mas em ambientes estruturais. Eventualmente, menos leal ao *media* e que cada vez vê menos televisão e menos dos mesmos programas. Atendendo às diferenças culturais observadas de país para país, entre outras, será necessário verificar se as tendências internacionais, também, são válidas a nível nacional, recorrendo a uma tipologia de dados muito mais adequada à sua análise, ou seja, às taxas de audiências condicionadas.

F.3.2. Objectivos e hipóteses

Neste capítulo procura-se avaliar a repercussão de determinados factores estruturais nos índices de lealdade programática, assim como a sua quantificação, considerando dados agregados e referentes ao universo televisivo. Estas análises permitem que seja avaliado o grau de regularidade da exposição, assim como contribuir para a selecção dos suportes a incluir em planos de *media*, em função do grau de duplicação de audiências observado. As análises realizadas, também, decorrem de posteriormente se pretender definir possíveis abordagens metodológicas recorrendo aos dados dos ficheiros de base de audimetria ou a dados agregados que permitam contribuir para a previsão dos *ratings* de séries programáticas, considerando os comportamentos dos telespectadores a este nível. Praticamente a totalidade dos trabalhos publicados identificados, apenas incluem na amostra em análise episódios sequenciais e incidem maioritariamente nas faixas horárias da noite.

Consequentemente, as análises a realizar incluem maiores desfasamentos temporais entre os diferentes episódios de um programa, transmitidos no decorrer de todo o dia nos principais canais generalistas. Este comportamento ainda que não seja, por si só, um factor explicativo da evolução de alguns dos indicadores de audimetria, ajuda-nos a compreender se a estabilidade que, por vezes, se observa nos índices de audiências se deve a que parte do painel desenvolva hábitos de consumo, apresentando lealdade programática.

Permite-nos tentar avaliar onde o painel de audimetria nacional se insere na dicotomia de selecção programática activa versus passiva. Os períodos horários, os canais e as tipologias programáticas em princípio devem-se diferenciar ao nível da evolução destes indicadores. No seguinte *Quadro F.3.2.1* estão representadas as hipóteses em avaliação neste apêndice.



Quadro F.3.2.1 - Hipóteses a testar quanto à repercussão de factores estruturais na duplicação de audiência

Uma vez que a maior parte do painel nacional não tem acesso á TV Cabo, ou os lares que o têm apresentam reduzidos índices de audiências para estes canais, e atendendo à regularidade que, por vezes, se observa nos valores dos *ratings*, seria de esperar razoáveis índices de duplicação, obviamente que dependentes dos períodos horários e das tipologias programáticas envolvidas. As variáveis em análise correspondem à frequência de transmissão de um programa, aos *ratings* dos programas, ao canal, à faixa horária e à tipologia programática, avaliando-se a interacção existente entre estes factores.

Face aos resultados das análises anteriores espera-se que estes índices sejam maximizados por parte dos canais privados no decorrer da maior parte da emissão diária. O canal RTP1 apenas deverá apresentar índices com uma maior dimensão para os espaços informativos e alguns dos seus programas sucessores. A popularidade de um programa deverá contribuir para o seu acréscimo, assim como uma maior frequência de transmissão semanal. As tipologias que apresentem uma sequência narrativa, como, por exemplo, as telenovelas ou as séries, não devem corresponder, necessariamente, às que maximizam os valores de duplicação de audiências. Os espaços informativos podem apresentar valores superiores ou no mínimo equiparáveis, uma vez que a bibliografia da especialidade associa-lhes em maior grau as motivações instrumentais dos telespectadores.

Relativamente à repercussão das faixas horárias, verificámos que os períodos mais penalizados ao nível da variabilidade inter-anual da variável *Tvr_Período%* corresponde à hora do almoço e à tarde dos dias laborais. As faixas horárias associadas a períodos de lazer apresentam uma maior estabilidade, de ano para ano. Considerando a variabilidade bimensal a maior parte dos programas da noite, também apresenta uma menor variabilidade de *ratings* ao fim-de-semana. Uma vez que as grelhas de programas, pelo menos dos canais privados, são estruturadas nos períodos pós-laboral por forma a incentivar a duplicação de audiências transmitindo-se programas na forma de episódios, por exemplo, as telenovelas e séries, aliado à maior estabilidade no decorrer dos anos das taxas médias de audiências para o Total da TV espera-se que a duplicação de audiências seja razoável e talvez superior ao fim-de-semana em horários pós laboral.

Outros períodos, como, por exemplo, as tardes de fim-de-semana ainda que correspondam a períodos de lazer devem apresentar índices muito inferiores, especialmente como consequência do pouco interesse que suscitam as demasiado frequentes retransmissões de filmes. No decorrer dos anos, as tardes e as manhãs dos dias laborais apresentam uma maior variabilidade nos valores das taxas médias de audiências para o total da TV, apesar da composição sociodemográfica da sua audiência estar em princípio associada em maior grau a quem não exerce uma actividade laboral. Consequentemente, estes índices não têm que ser necessariamente muito superiores nestes períodos horários.

F.3.3. As bases de dados e metodologias

Para a análise dos índices de lealdade programática vamos utilizar os índices de duplicação das audiências calculados através das análises individuais do *software* Telereport que cruzam os vários episódios dos programas seleccionados e identificados segundo o dia e hora a que correspondem. O *software* proporciona as suas taxas de audiências condicionadas que representam a percentagem do alvo que seleccionou o programa em linha (Prog_I) e que, também seleccionou o programa em coluna (Prog_J). Podemos seleccionar o nível programático global que considera a duração programática total ou seleccionar o nível parcial, isto é, cada uma das partes que integram os programas em análise. Com base no *software* Telereport podemos construir quadros onde consta a informação que se exemplifica no seguinte *Quadro F.3.3.1*.

	Prog 1 - Dia 1	...	...	Prog 1 - Dia n
Prog 1 - Dia 1	$P(\text{Prog 1 - Dia 1} / \text{Prog 1 - Dia 1})$	...	...	$P(\text{Prog 1 - Dia 1} / \text{Prog 1 - Dia n})$
Prog 1 - Dia 2	$P(\text{Prog 1 - Dia 2} / \text{Prog 1 - Dia 1})$	...	...	$P(\text{Prog 1 - Dia 2} / \text{Prog 1 - Dia n})$
...	...	...	...	...
Prog 1 - Dia n	$P(\text{Prog 1 - Dia n} / \text{Prog 1 - Dia 1})$	...	...	$P(\text{Prog 1 - Dia n} / \text{Prog 1 - Dia n})$

Quadro F.3.3.1 - Índices de duplicação das audiências de lealdade programática

Consideram-se nas análises os programas dos três principais canais generalistas, ou seja, da SIC, TVI e RTP1. Nas análises individuais do *software* Telereport apenas consta a identificação do programa, dia e hora de início da sua transmissão. Não é possível através do *software* agregar a esta base de dados as restantes variáveis de caracterização da grelha de programação. Consequentemente, a partir dos quadros descritos anteriormente (*Quadro F.3.3.1*) e construídos para cada um dos programas em análise será necessário construir uma segunda base de dados onde na primeira coluna consta o programa de referência e na segunda coluna consta o programa subsequente. Adiciona-se o índice de duplicação de audiências e algumas variáveis de caracterização da programação, entre as quais as seguintes: o dia da semana ou diferenciado em semana vs. fim-de-semana, a hora de início, a duração, a tipologia programática de grau 1, 2 e 3, o canal, a faixa horária e o número de transmissões por semana (*Quadro F.3.3.2*).

Programa Referência	Programa Subsequente	Índices de Lealdade Programática	Variáveis de Caracterização da Programação
Espaço Informativo da noite no Dia X	Espaço Informativo da noite no Dia X+1	% da audiência que viu pelo menos 50% do programa de Ref. (Por Ex. Espaço Informativo da noite no Dia X) que, também, vê pelo menos 50% do Programa Subsequente (Por Ex. Espaço Informativo da noite no Dia X+1)	...
Espaço Informativo da noite no Dia X	Espaço Informativo da noite no Dia X+2	...	...
...	...	...	...
Espaço Informativo da noite no Dia X	Espaço Informativo da noite no Dia X+n	% da audiência que viu pelo menos 50% do programa de ref. (Por Ex. Telejornal Dia X) que, também, vê pelo menos 50% do programa subsequente (Por Ex. Telejornal Dia X+n)	...

Quadro F.3.3.2 – Bases de dados para a análise de índices de lealdade programática

De acordo com Zubayr (1999), as bases de audimetria são as mais adequadas para a análise destes comportamentos estruturais de consumo televisivo. As análises que recorrem a inquéritos apresentam grandes limitações e incorrecções, uma vez que dependem, entre outros factores, da capacidade de memorização dos telespectadores:

Webster and Litchy (1992) report evidence that many diary keepers do not report their actual viewing behaviour, but rather their remembered behaviour at the end of the diary week. *Brosius, Wober and Weimann (1992) and Twyman (1983)* believe that recording watched programs from memory a few days later puts single broadcast programs at a disadvantage and overemphasizes habits, which again raise repeat viewing rates. Third, the measurement technique itself could be responsible for deviation in results. While people-meters record every second of viewing behaviour, diary keepers are usually asked to note their viewing behaviour in 15 minutes intervals.

Em *Zubayr (1999)*, a variável dependente corresponde às taxas de audiências conjuntas e nas análises a realizar utilizam-se as taxas condicionadas, assim como um maior horizonte temporal. Segundo esta autora, a maioria dos autores não considera nos seus trabalhos indicadores para cada um dos programas, mas, sim, uma única distribuição de frequências para todos os episódios de um programa transmitido em determinado período de tempo: “studies unfortunately had just a single distribution for all episodes of a program in a given time period”. Em *Sherman (1995)*, também, se utiliza essa abordagem: “Barwise (1986) translation which is used in much of the previous research in repeat viewing was used to produce a summary statistic of repeat viewing for each series. ... it imputes an overall rate of repeat viewing rate of repeat viewing for multiple episodes.”

A tipologia de dados utilizada na presente tese corresponde à mais adequada para a caracterização destes comportamentos. Em *Webster (2006)*, consta uma crítica à utilização das audiências conjuntas para a avaliação da duplicação de audiências entre programas adjacentes, o mesmo se verificando com os índices de lealdade programática:

For those who want to understand how program content and structures affect the way audiences flow through the evening, or any day part, primary duplication is a much more satisfying measure. ... the duplicated audience is most appropriate for analysis concerned with the absolute audience sizes. For example, it is useful in building models that predict reach and frequency.

Outra questão a definir corresponde à duração da exposição a considerar nas análises e que não consta em *Zubayr (1999)*, assim como na maioria das publicações existentes, levando a pressupor que um segundo de exposição televisiva a cada programa permite ao telespectador ser contabilizado para o cálculo dos índices de “repeat viewing”.

Deve ser considerado, um horizonte temporal mais alargado denotando algum envolvimento dos telespectadores com os conteúdos transmitidos. Nas análises a realizar definiu-se dois cenários, um dos quais considerando a duração que impera na bibliografia da especialidade e que corresponde a um segundo e outro cenário com uma percentagem mínima de exposição que corresponde a 50% da duração do suporte em análise. Os resultados apresentados neste capítulo correspondem ao segundo cenário que foi considerado. Portanto, consideram-se os telespectadores que viram pelo menos metade da duração de cada um dos programas, ao contrário da literatura da especialidade que em geral pondera qualquer telespectador independentemente da duração do contacto.

Hoje em dia, com a existência do controle remoto facilmente alguém poderá estar exposto a um programa durante alguns segundos, não devendo ser realmente considerados na duplicação de audiências, que pretende caracterizar e analisar comportamentos de lealdade programática e não uma exposição meramente pontual e pouco representativa do seu comportamento. Adoptou-se metade da duração da programação, uma vez que para a definição da contra-programação de um programa, por definição nas agências de planificação de meios, também se considera metade da duração dos programas como valor de referência. As unidades estatísticas correspondem a pares de episódios e transmissões de programas.

Para serem incluídos na amostra, os programas têm que apresentar para o período em análise pelos menos duas transmissões. Associou-se a cada par de programas um conjunto de variáveis explicativas, sendo a variável dependente considerada o valor do índice de duplicação de audiências quantificado através das taxas de audiências condicionadas, ou seja, Taxa ($\text{ver episódio } j / \text{ver episódio } i$) que representa a percentagem dos telespectadores que viram pelo menos 50% do programa i que também viram pelo menos 50% do programa j . As variáveis consideradas e hipóteses a testar constam do *Quadro F.3.2.1*. Para a avaliação da repercussão de variáveis estruturais nos índices de duplicação recorreu-se a análises Anova a I e II factores ou a testes T para médias quando se consideram variáveis explicativas nominais e binárias, ou aos equivalentes algoritmos não paramétricos e aos coeficientes de correlação linear de pearson quando se consideram variáveis explicativas numéricas.

F.3.5. Resultados

I. Avaliação das Influências Estruturais na Evolução dos Índices de Lealdade Programática

Calculando os coeficientes de correlação linear de pearson entre os índices de lealdade programática e determinadas variáveis explicativas estruturais quantitativas ou realizando análises Anova a um e dois factores quando as variáveis independentes são nominais foi possível retirar as seguintes conclusões:

➤ Existe associação entre os índices de duplicação e as tipologias programáticas definidas nas bases de dados Telereport (Análise Anova, p valor $<0,000$).

Os espaços destinados à transmissão de filmes apresentam os menores valores, não excedendo em média os 25% para o período em análise. As restantes tipologias, como, por exemplo, os espaços informativos, as telenovelas e as séries apresentam índices ligeiramente superiores. As telenovelas pelo seu carácter narrativo correspondem aos programas que maximizam este indicador, apresentando valores no intervalo [40%, 55%] e os espaços informativos com valores no intervalo [30%, 40%].

➤ Não existe associação entre os índices de duplicação e os canais generalistas.

Comparando os diversos canais generalistas não se observam diferenças estatisticamente significativas nos valores médios dos índices de duplicação considerando a totalidade da emissão diária.

Em função dos resultados das análises realizadas aos valores de *ratings* e aos valores das taxas médias de audiências para o total da TV, verificámos que talvez pudéssemos associar aos canais privados em maior grau as motivações ritualísticas dos telespectadores que integram a sua audiência e ao canal estatal as motivações instrumentais, isto é, mais intencionais. Esta realidade poderia, eventualmente, conduzir a valores diferenciados de duplicação de audiências entre os canais públicos e os canais privados, inclusive considerando a totalidade dos dados.

Porém, as motivações ritualistas não implicam necessariamente uma exposição mais regular. O conceito de hábito, como já foi salientado, não implica a existência de uma forte regularidade na exposição dos telespectadores, mas, sim, uma maior probabilidade e tendência para seleccionar determinados suportes no decorrer do tempo.

Os telespectadores que apresentem motivações instrumentais procuram determinados conteúdos, como, por exemplo, no canal estatal RTP1 onde existe claramente uma procura dos conteúdos informativos. Logo, nestes contextos a duplicação das audiências resultará em maior grau de eventuais “preferências” programáticas e por determinadas tipologias de conteúdos, estando a regularidade da exposição condicionada por parte de factores estruturais e individuais.

Verificámos que de acordo com a evolução inter-anual e bimensal dos valores de ratings e das taxas médias de audiências para o total da TV, o canal estatal RTP1 corresponde ao que consegue estabilizar em maior grau os seus índices de audiências, desde que excluindo os programas mais populares.

No entanto, as alterações horárias de alguns programas, também, incutem variabilidade nos indicadores de audimetria. Apesar da procura do seu público por determinados conteúdos, impera a influência dos seus factores de disponibilidade, entre outros. Quando imperam as motivações ritualísticas dos telespectadores, a duplicação da audiência entre os suportes resulta de comportamentos mais passivos, orientados para o *media* e em menor grau para os conteúdos programáticos transmitidos. Neste contexto, os telespectadores assistem sucessivamente ao mesmo programa, apenas porque o televisor está ligado, dia após dia, nesse período horário e em geral é realizada uma selecção inicial ao nível do canal, em função da grelha de programas com que cada telespectador mais de identifica (lealdade ao canal).

Verificámos que as alterações horárias nos canais privados podem provocar uma grande variabilidade ao nível da evolução dos ratings, inclusive considerando os programas com um maior exposição ao painel, o que permitir confirmar a existência de uma exposição direccionada em maior grau ao *media*.

Observámos que ao fim-de-semana as alterações horárias originam uma menor variabilidade dos valores de ratings, o que se poderá atribuir a uma maior disponibilidade do painel para ver TV ou ao envolvimento e

interesse despertado pelos conteúdos programáticos.

No próximo capítulo constam as análises realizadas aos efeitos de herança e aos fluxos de audiências entre os canais, por forma a confirmar estes resultados. Nomeadamente, avaliar o grau de passividade do painel e o grau de independência existente entre os públicos de cada canal.

Portanto, independentemente da motivação que maioritariamente impera no público de cada um dos canais, a ordem de grandeza dos índices de duplicação dos vários canais generalistas é equiparável a nível global.

Em princípio poderíamos pensar que a procura por determinados conteúdos levaria a uma maior duplicação de audiências, mas os factores de disponibilidade, também, condicionam os resultados. Quando as motivações são ritualísticas e se desenvolvem hábitos de consumo, a exposição também se encontra condicionada pelo padrão de tempos vazios e eventualmente de lazer dos telespectadores, sendo especialmente difícil no primeiro contexto que se verifique alguma regularidade no consumo televisivo. A exposição televisiva encontra-se influenciada pelas existência ou inexistência de rotinas laborais e sobre as quais grande parte dos telespectadores, frequentemente, não têm poder de decisão, assim como pelo grupo em que este se encontra inserido.

➤ Verificamos que existem interacções entre os períodos horários e as tipologias programáticas.

Em relação aos filmes, verificamos que os programas que são transmitidos no decorrer da noite apresentam índices ligeiramente superiores aos da tarde. As telenovelas transmitidas após o espaço informativo da noite, também superam ligeiramente as telenovelas do pré prime time. Verificámos anteriormente que a variável *Tvr_Período%*, também apresenta uma maior estabilidade nos períodos pós-laboral considerando a variabilidade inter-anual, mas no período bimensal os menores valores de variabilidade estão associados às faixas horárias laborais. Nos períodos de maiores audiências existe uma maior estabilidade ao fim-de-semana.

No entanto, independentemente do grau de variabilidade das taxas médias de audiências para o total da TV, quanto maior for a sua dimensão maior será a probabilidade de existir a duplicação de audiências. Um programa ao qual assistem 100 telespectadores mais facilmente poderá apresentar 50 telespectadores no episódio seguinte do que um programa que apresenta uma exposição a apenas 60 indivíduos. Logo, facilmente se compreende que as telenovelas do horário nobre maximizam os valores destes índices relativamente às telenovelas transmitidas no pré-prime time.

No decorrer da noite, as telenovelas apresentam uma tendência decrescente quanto aos seus valores de duplicação das audiências. Uma vez que se reduzem os *ratings*, os programas que apresentam um menor número de telespectadores, também apresentam uma menor probabilidade de os manterem de episódio para episódio.

Não se verifica a existência de uma audiência de menor dimensão, mas mais leal. Em relação aos filmes verifica-se o inverso. Os filmes da noite maximizam estes índices, apesar dos seus menores valores de *ratings*, indiciando a existência de uma audiência de menor dimensão, mas mais leal aos conteúdos. Uma vez que os filmes mais interessantes são transmitidos em horários mais tardios e grande parte do painel não apresenta TV Cabo, existindo uma “preferência” por esta tipologia de programas, faz com que este segmento assista sempre que possível aos filmes que considera mais interessantes ou que estão disponíveis em determinado momento. Os filmes transmitidos no decorrer da tarde ao fim-de-semana correspondem, frequentemente, a retransmissões dos mesmos.

Nestas faixas horárias mais facilmente grande parte do painel poderá optar por outras actividades de preenchimento do seu tempo de lazer. Consequentemente, a duplicação de audiências será menor.

Considerando os espaços informativos, o programa da tarde do canal RTP1 apresenta maiores valores de duplicação relativamente ao programa da noite. Para o canal TVI verifica-se o inverso, enquanto que a SIC apresenta para ambos os programas valores equiparáveis, atendendo ao período em análise.

Os valores de *ratings* do espaço informativo da tarde do canal RTP1 são mais estáveis em relação aos valores do espaço informativo da noite, e, logo, uma maior duplicação associa-se a uma maior estabilidade de valores de *ratings*.

Porém, apesar de ambos os espaços informativos do canal SIC apresentarem a mesma tendência ao nível dos valores de duplicação das audiências, os seus programas diferenciam-se ao nível da variabilidade dos correspondentes valores de *ratings*. Logo, devemos salientar que a ordem de grandeza da duplicação de audiência entre os diversos episódios de um programa, nem sempre permite avaliar qual será a evolução ao nível da variabilidade dos seus *ratings*. Consequentemente, a previsão dos indicadores de audimetria por análise das selecções históricas do público que apresenta uma maior regularidade na sua exposição apresenta-se como insuficiente e condiciona as abordagens mais adequadas para a construção dos planos de *media*.

➤ Inter-relacionando as variáveis canal, tipologias e faixas horárias foi possível retirar as conclusões que constam dos seguintes parágrafos.

O canal RTP1 apresenta maiores valores de duplicação para o espaço informativo da tarde, uma vez que apresenta uma maior estabilidade horária e ao nível da sua contra-programação e das taxas médias de audiências para o total da TV nos períodos horários em que é transmitido, aliado às motivações instrumentais do seu público, ou seja, à procura de informação. Neste canal, os espaços informativos da noite com as frequentes alterações horárias a que estão sujeitos e ao nível da maior variabilidade das suas taxas médias de audiências para o total da TV (ainda que muito superiores em dimensão face às das faixas horárias da hora do almoço) apresentam menores valores de duplicações das audiências.

A existência de uma audiência total superior não implica necessariamente que se observem maiores valores de duplicação das audiências.

O canal SIC uma vez que apresenta uma posição mais consolidada para o período em análise, no âmbito dos canais privados, consegue equiparar a duplicação de audiências de ambos os espaços informativos.

O canal TVI uma vez que em alguns dos anos considerados apresenta uma tendência ascendente no valor do seu *share* médio, o que foi alcançado retirando telespectadores ao canal SIC e à RTP1, consegue maximizar a duplicação de audiências nas faixas horárias da noite. Isto também poderá ser atribuído à popularidade dos seus programas de ficção nacionais que cada vez mais suscitam o interesse do público e acabam por influenciar por efeito de lead out os seus espaços informativos, incutindo-lhes uma maior duplicação de audiências, face ao programa equivalente da tarde. Os espaços informativos da tarde dos canais concorrentes, uma vez que apresentam uma posição mais consolidada superam a TVI a este nível, para o período em análise.

Em *Fernandes (2001)*, com base na análise documental realizada, apenas é referido que podemos observar o denominado “*double jeopardy effect*”, isto é, quando uma estação de televisão apresenta uma fraca audiência e telespectadores pouco leais e irregulares. Verificamos que, nem sempre isto se verifica, por exemplo, no âmbito dos filmes.

Em relação à sua constatação de que a duplicação da audiência nos programas emitidos em day time, durante a semana poder ser superior ao valores da noite, também verificamos que nem sempre se verifica, uma vez que os resultados estão fortemente condicionados pelo canal e por parte das tipologias programáticas envolvidas.

➤ Existe uma reduzida correlação, mas positiva entre o desempenho de um programa e os índices de lealdade programática ($\rho=0,35$).

Obviamente que o desempenho de um programa associa-se com maiores valores dos índices de duplicação, como, por exemplo, se verifica com as telenovelas mais populares do início da noite, ainda que a relação poderá não ser linear, o que leva a um menor valor do coeficiente de correlação.

No entanto, as alterações horárias dos programas que apresentam um maior desempenho podem contribuir para uma significativa redução dos valores de duplicação de audiências. Verificamos que ao fim-de-semana é mais fácil observar menores intervalos de variabilidade de valores de *ratings* decorrente das alterações horárias dos programas, para além de nos períodos de maiores audiências ao fim-de-semana existir uma maior estabilidade ao nível das taxas médias de audiências para o total da TV.

➤ Existe alguma correlação linear entre a frequência de transmissão de um programa e os índices de duplicação ($\rho=0,60$). A cada programa foi associado o número de transmissões semanais e os programas que apresentam maior número de transmissões

maximizam estes índices. Estes programas acabam por estar associados às telenovelas, logo inevitavelmente os índices de duplicação são superiores.

II. Análises aos índices de fidelidade

O índice de fidelidade mede a relação existente entre a audiência média e a audiência total do programa, ou, também, a proporção de tempo médio gasto pela audiência na duração global do programa.

Na caracterização do processo de selecção programática é relevante identificar quais são os programas seleccionados com uma maior regularidade por parte de cada um dos segmentos, como também caracterizar o seu grau de permanência.

Neste item pretendemos identificar quais são as tipologias programáticas que maximizam este indicador para cada um dos segmentos sociodemográficos representados nas bases de dados e por canal. Associaram-se intervalos de variabilidade desta variável a determinados programas. A variável fidelidade apresenta uma grande amplitude de valores considerando o universo televisivo e a totalidade dos programas.

IIa. Alguns Resultados para o Canal SIC

Por aplicação de algoritmos de *clustering* simbólicos aos valores de fidelidade dos segmentos associados às variáveis classes sociais, sexo e faixas etárias, os agrupamentos obtidos foram os seguintes:

□ O primeiro agrupamento minimiza os valores de fidelidade para todos os segmentos representados nas bases de dados Telereport. Inclui a maior parte dos filmes da noite e da tarde, os programas matinais, sejam estes *talk shows* ou os juvenis, alguns eventos, por exemplo, os Globos de Ouro e os programas do final da noite. O que parece influenciar os reduzidos valores de fidelidade que estes programas apresentam poderá ser o seu horário de transmissão e, também, por corresponderem a espaços televisivos dedicados à transmissão de filmes, uma vez que inúmeras vezes consistem em repetições.

□ O segundo agrupamento corresponde ao segundo conjunto de programas com maiores valores de índices de fidelidade para todos os segmentos considerados. Inclui programas da tipologia telenovelas ou séries, repartidos ao longo do dia e também o espaço informativo da tarde. Estes programas apresentam uma grande diversidade de horários de transmissão, incluindo também programas do final da noite e da tarde, correspondendo alguns deles inclusive a retransmissões. Alguns programas beneficiam dos efeitos de herança, como, por exemplo, o programa que sucede o espaço informativo da tarde, a telenovela que antecede o espaço informativo da noite ou o programa de humor que o sucede. As telenovelas e as séries pelo seu próprio carácter narrativo e sequencial favorecem maiores valores dos índices de fidelidade das audiências.

□ O quinto agrupamento apresenta os maiores valores dos índices de fidelidade para todos os segmentos.

Corresponde a programas de muito curta duração e aos programas que são transmitidos nas faixas horárias que intercalam o espaço informativo da noite e a principal telenovela da noite ou aos programas transmitidos no período designado de horário nobre.

IIb. Alguns Resultados para o canal TVI

Realizando idênticas análises para o canal TVI foi possível retirar as seguintes conclusões:

□ No primeiro agrupamento constam programas essencialmente do fim-de-semana e das tipologias de filme ou séries. Correspondem a filmes transmitidos no decorrer da tarde e da madrugada ao fim-de-semana e, também, no decorrer da noite em dias laborais. Ao longo da semana, também, estão representados os talk-shows da tarde e da noite e o espaço informativo da noite. Estes programas minimizam os valores de fidelidade para a classe social AB.

□ O segundo agrupamento apresenta os índices mais elevados para todos os segmentos. Inclui programas de duração muito reduzida transmitidos entre programas com alguma popularidade, assim como o programa juvenil das tardes e os filmes da noite em determinados dias.

□ O terceiro agrupamento corresponde ao segundo conjunto de programas com os maiores valores de fidelidade para a maioria dos segmentos, mas excluindo a classe social AB. Inclui programas matinais de fim-de-semana destinados às crianças, filmes da noite, o concurso do final tarde e as telenovelas dos dias laborais transmitidas antes do espaço informativo da noite ou no período da madrugada.

Os filmes transmitidos à noite, ao contrário do que se verifica na SIC, conseguem apresentar razoáveis valores de fidelidade.

□ O quarto agrupamento corresponde ao que apresenta menores índices de fidelidade para todos os segmentos considerados nas análises.

Inclui filmes de fim-de-semana e do fim da tarde, os programas matinais em dias laborais, programas com um único episódio, alguns dos quais repetidos, como, por exemplo, algumas peças de teatro, assim como os concertos, os jogos de futebol, as telenovelas ou galas.

No horizonte temporal em que este canal ainda apresenta uma tendencia crescente dos valores de shares de audiências, observamos que no período matinal não consegue manter os telespectadores no seu decorrer, contrastando com o canal SIC no decorrer deste período.

IIc. Alguns Resultados para o Canal RTP1

Para o canal estatal RTP1 obtiveram-se os seguintes agrupamentos:

□ O primeiro agrupamento corresponde ao segundo grupo que apresenta os maiores valores dos índices de fidelidade para todos os segmentos considerados nas análises. Inclui o espaço informativo da tarde no decorrer de toda a semana e apenas o espaço informativo da noite dos dias laborais, o programa de humor após o espaço informativo da noite, o concurso que antecede o espaço informativo da noite (“O Preço Certo”) e o concurso do horário nobre após o espaço informativo da noite, a telenovela após o espaço informativo da tarde e as séries de humor de fim-de-semana antes do espaço informativo da tarde.

Verificamos que correspondem aos dois espaços informativos e a programas que beneficiam dos efeitos de herança e que são transmitidos antes ou após os espaços informativos.

□ O segundo agrupamento é o que apresenta os valores mais reduzidos dos índices de fidelidade. Inclui filmes transmitidos ao longo de toda a semana e no decorrer da noite e da tarde; o programa de debate da noite “Prós e Contras”; os programas matinais semanais; o concurso da noite e mini-séries do final da noite.

□ O terceiro agrupamento representa o segundo conjunto de programas com os valores de índices de fidelidade mais reduzidos. Corresponde, essencialmente, a programas juvenis e matinais de fim-de-semana, aos filmes do final da noite durante a semana, a programas de informação desportiva e ao *talk show* da tarde transmitido de Segunda a Sexta.

□ O quarto agrupamento é o terceiro conjunto de programas com os valores de índices de fidelidade mais elevados. Inclui programas de desenhos animados, programas de Economia e Política, alguns concursos, programas desportivos e o espaço informativo da noite de fim-de-semana.

F.4. ANÁLISE DE ÍNDICES DE DUPLICAÇÃO DAS AUDIÊNCIAS ENTRE PROGRAMAS E FAIXAS HORÁRIAS

F.4.1. Introdução

Este capítulo visa a análise dos índices de duplicação de audiências entre programas e períodos horários próximos ou adjacentes e dos fluxos de audiências entre canais, uma vez que estes comportamentos de consumo televisivo podem condicionar a evolução dos principais índices de audiências. Tendo como objectivo contribuir para a previsão de valores de *ratings* ou a identificação de métodos alternativos para a construção de planos de *media*, a desagregação das análises a este nível poderá justificar-se para os períodos horários que maximizam a variabilidade observada nos principais indicadores de audimetria. Os efeitos de herança entre programas adjacentes apresentam uma grande relevância para o planeamento de *media* e associam-se com os factores estruturais e com a previsão das audiências (*Fernandes, 2001*):

Dos factores que estão associados às audiências, os estruturais referem-se às audiências potenciais (previsão) e à audiência disponível ... No que concerne aos factores associados aos media também se distinguem determinantes estruturais e individuais. No primeiro grupo, encontra-se ... a duplicação de audiências que se relaciona com o facto de dois programas diferentes terem uma audiência comum (quando os programas são emitidos em períodos adjacentes também conhecido por “efeito de herança”).

Na realidade os efeitos de herança não dizem apenas respeito a programas adjacentes, apesar desta tipologia de análise imperar na bibliografia da especialidade. Os efeitos de herança, também, não, representam a duplicação de audiências entre os programas, mas, sim, a duplicação que resulta da passividade do público. Contudo, alguns autores (*Webster, 2006*) salientam a dificuldade de isolar os diferentes tipos de lealdade que contribuem para a duplicação das audiências entre os programas de determinado canal e dia:

Inheritance effects are one of the most important of all television audience behaviours ... neither the growth of alternative delivery systems nor the universal penetration of remotes seem to have diminished the phenomenon. Inheritance effects are to this day the foundation of programming strategies that have been in use for several decades (...).

Yet many studies that claim to document inheritance effects are hampered by their inability to directly observe the duplicated audience. Further, there is a lack of conceptual clarity on how to measure inheritance effects that leads to widely divergent, even contradictory results ... When the results on primary duplication are considered, there is certainly much variance left to be explained ... For those who want to understand how structures affect the way audiences flow through the evening or any day part, primary duplication is a much more satisfying measure... Audience flow is a topic of enduring interest to broadcasters ... Legendary TV programmers built their reputation in part, on their ability to manage audience flow ... The current study replicates Webster (1985) using a national sample, thus offering a look at the same relationships 22 years later. ... Whether inheritance effects will continue to hold up in the future is an open question. Digital video recorders and video-on-demand are at this writing on verge of widespread adoption. If that happens and if viewers use them to aggressively reconfigure what the networks are programming then the audience flow, as it is known today may well become a thing of the past. But those are very big “ifs”. In large theoretical terms, this research should serve as a reminder that television program choice is the complicated result of many factors.

Para além da repercussão dos efeitos de herança na evolução dos índices de audiências, a exposição aos canais concorrentes também deve ser avaliada e na literatura da especialidade estes comportamentos denominam-se de “zapping”. A audimetria mede estes comportamentos de consumo televisivo com uma maior fiabilidade, uma vez que caracteriza a exposição ao segundo (*Fernandes, 2001*):

O zapping conduz a que a medição da audiência, tenha de se adaptar a esta nova realidade e a um grau elevado de sofisticação. ... A audiplan é o novo serviço da Marktest Audimetria e permite a medição ao segundo, bem como uma análise detalhada dos telespectadores com a mensagem publicitária, captando cerca de 50 a 60 mudanças de canal por dia. Este sistema faculta uma avaliação do zapping entre canais.

Ao contrário do que *Fernandes (2001)* refere a evolução dos “ratings” não pressupõe a existência de regularidade dos comportamentos de consumo televisivo, sendo a audimetria o sistema de informação que viabilizou constatar internacionalmente esta realidade para grandes horizontes temporais:

Ora, o discurso dos ratings refere-se aos programas vistos, mas o que poderá vir a ser um sucesso, não é controlado, pois existe a crença de que o comportamento das audiências se mantém sem grande variação.

As referências que constam no texto apresentado neste âmbito referem-se apenas a trabalhos internacionais. Este capítulo pretende, não só, a análise dos índices de duplicação entre programas num mesmo canal, mas, também, dos fluxos de audiências entre canais recorrendo essencialmente a dados de duplicação primária, isto é, a taxas de audiências condicionadas que correspondem à tipologia de dados realmente adequada à sua análise e menos utilizada. A nível nacional não foi identificado qualquer trabalho nesta área, dificuldade que aparentemente outros autores também demonstram ter tido, uma vez que, por exemplo, em *Fernandes (2001)* não se encontra nenhuma referência nacional. Inclusive a nível internacional existe uma lacuna bibliográfica quanto à análise dos fluxos de audiências entre canais ou entre programas que não sejam adjacentes e nenhum trabalho foi identificado ao no âmbito das faixas horárias. Os efeitos de herança são essencialmente avaliados considerando apenas o fluxo existente a partir de determinado programa para o subsequente e, também, não têm em consideração o recíproco, avaliando a duplicação de audiências face aos programas precedentes que, também, foi considerado nesta tese. Ao contrário da maioria dos trabalhos publicados neste âmbito as análises não se debruçam exclusivamente nos períodos de maiores audiências e recorre-se a uma tipologia de dados utilizada com uma menor frequência na literatura internacional da especialidade. Diferentes horizontes temporais podem apresentar uma evolução diferente destes indicadores (*Fernandes, 2001*):

A cada período horário correspondem estratégias diferentes e tipos de programação, face ao reconhecimento da existência de disparidade nos empregos de tempo por parte dos indivíduos, cujas características sociográficas também são diferentes.

Do ponto de vista metodológico contrastando com as publicações existentes recorre-se e sugerem-se abordagens metodológicas mais adequadas na área do *Data Mining*, em vez das generalizadas análises de Regressão Linear que representam a quase totalidade das publicações existentes. As análises a realizar contribuem para clarificar inúmeras incógnitas quanto ao grau de regularidade da exposição televisiva e sobre a sua repercussão na evolução dos índices de audiências. Os denominados efeitos de herança traduzem a tendência que se verifica nos telespectadores para permanecer sintonizados ao mesmo canal após a conclusão de um programa. Por um efeito de inércia os telespectadores são transportados para os programas subsequentes, condicionando os valores dos seus índices de audiências.

A quantificação, caracterização e a avaliação destes comportamentos tem suscitado o interesse do meio televisivo, ao longo das últimas décadas, pela sua repercussão na evolução dos índices de audiência, persistindo na actualidade, como consta dos parágrafos introdutórios.

Segundo *Webster (2006)*, os efeitos de herança são desde há longa data os pilares do desenvolvimento das estratégias de programação. Inúmeros trabalhos que visam a análise destes comportamentos encontram-se penalizados pela sua inaptidão para conseguir observar de forma directa a duplicação de audiências. Não existe uma grande clareza quanto à forma mais correcta de medir os efeitos de herança o que tem levado a resultados frequentemente contraditórios e divergentes, por exemplo, nos Estados Unidos. Constatação que *Fernandes (2001)*, também refere no âmbito das análises realizadas aos índices de “repeat viewing”.

As estações televisivas conscientes de que este fenómeno pode ser determinante para o seu bom desempenho, têm desenvolvido estratégias de planeamento que o favoreçam. Estas estratégias em geral traduzem-se pela transmissão de programas de maior popularidade em períodos adjacentes a programas menos apelativos incentivando à permanência do público por um maior horizonte temporal.

Estes comportamentos recorrentes são melhor entendidos quando enquadrados num marco teórico que reconhece os padrões de disponibilidade, as preferências e as inúmeras variáveis estruturais que restringem essas preferências (*Rust & Echambadi, 1989*). Ao longo das últimas décadas os índices de duplicação têm sido atribuídos a inúmeros factores. Segundo alguns autores resultam da lealdade que é dirigida a tipologias e aos canais, ou seja, da tendência do público para apresentar uma certa consistência nas suas preferências por determinado canal face aos restantes. Outros autores atribuem-no essencialmente a fenómenos de lealdade programática e aos efeitos de herança, uma vez que os índices obtidos excedem em dimensão os que resultariam de uma mera lealdade a canais. O conceito de efeitos de herança, traduz os índices de duplicação das audiências entre programas adjacentes quando estes são essencialmente resultado da inércia do público, ainda que em parte influenciados por comportamentos de lealdade.

Assim sendo, estruturando a grelha de programas por forma a incentivar a permanência da audiência através de programas consecutivos de um mesmo canal, poder-se-á conseguir um planeamento com sucesso (*Boemer, 1987; Cooper, 1993; Davis & Walker, 1990; Walker, 1988; Webster, 1985*). Segundo *Cooper (1993)* e *Webster (1985)* os programas herdaram por inércia audiência dos programas que os antecedem (efeito designado de *Lead_in*) ou sucedem (efeito designado de *Lead_out*). O efeito de *lead_in* é considerado por muitos autores como sendo mais relevante, ou seja, os programas beneficiam essencialmente de um maior desempenho do programa que os precede.

A aderência da indústria e dos académicos ao poder dos efeitos de herança baseia-se no pressuposto implícito de que o público é de certo modo passivo, assistindo a um programa simplesmente porque sucede outro que já estão a assistir. Independentemente da veracidade desta afirmação, programas adjacentes partilham uma elevada percentagem da audiência e um grande número de pessoas, também, tende a assistir à televisão em períodos adjacentes (*Webster, 1985*), logo a correlação tem que ser elevada. Segundo *Webster*, a duplicação de audiências entre programas adjacentes resulta em primeiro lugar de factores de disponibilidade do público.

Os efeitos de herança podem ser encarados como um caso especial de duplicação de audiências num mesmo canal. Os índices associados à lealdade a canais têm sido utilizados para justificar que programas adjacentes ou quase adjacentes apresentam um caso especial de lealdade que se denomina de efeitos de herança. *Kirsh e Banks (1962)* foram os primeiros a reparar na existência de elevados níveis de duplicação de audiências entre programas adjacentes, seguindo-se os trabalhos realizados por *Goodhardt, Ehrenberg e Collins (1975)*.

Desde então são inúmeros os trabalhos que visam documentar os efeitos de herança e a maioria dos quais pode ser agregado em três categorias segundo a tipologia de dados utilizada nas análises. A primeira categoria de trabalhos tem em consideração a “*duplication of viewing law*” (*Goodhardt, Ehrenberg & Collins, 1975*) como ponto de partida e adaptam-na ao estudo dos efeitos de herança, recorrendo a dados de duplicação totais.

A segunda e maior fracção dos trabalhos limitam-se a comparar as correlações de *lead_in* e de *lead_out* ou de valores de *shares* programáticos sob vários contextos. A terceira categoria minoritariamente abordada, assim como a primeira utilizam evidências directas da existência de efeitos de herança, mas consideram como variável dependente uma variável distinta representada pela duplicação primária.

Entre as técnicas analíticas mais utilizadas ao nível da análise dos índices de duplicações a mais directa e uma das iniciais foi desenvolvida por *Goodhart (1966)*, alargando este conceito a programas transmitidos em dias diferentes. Segundo *Goodhart*, a percentagem de telespectadores que vê dois programas do mesmo canal em diferentes dias ou no mesmo dia (*rsk*) é função dos seus *ratings* (*rs* e *rk*) e de uma constante dividida por 100 representada pela equação $rsk = rs * rk * k / 100$. Sob esta formulação quando *k* assume o valor unitário existe independência estocástica entre as audiências programáticas. Posteriormente, esta abordagem foi criticada por *Headen, Klompmaker e Rust (1979)* e *Webster (1985)*. *Webster* desenvolveu um modelo baseado na RLM que incluía como variáveis explicativas, as seguintes: o *rating* de *lead_in*, o *rating* de *lead_out*, as sinergias tipológicas e o número de programas que se iniciam em contra-programação quando conclui o primeiro programa inserido em cada par. Estas quatro variáveis explicam 85% da variância da duplicação de audiências (*rst*) e a variável *rst* representa a percentagem da audiência total que assiste a ambos os programas ($p(i \wedge j)$).

O trabalho realizado por *Webster (2006)*, nos USA, consiste em uma réplica de um trabalho anteriormente desenvolvido pelo autor e que permite verificar se as mesmas relações se mantêm após 22 anos. Recorre às três tipologias de dados enumeradas anteriormente. Segundo *Webster* o primeiro conjunto de trabalhos apresenta a vantagem de oferecer uma medida directa deste comportamento, sabendo-se qual a percentagem da audiência que viu ambos os programas ($p(i \wedge j)$).

No entanto, estes modelos carecem de alguma qualidade, por exemplo, uma vez que programas com reduzidos níveis de exposição vão certamente apresentar reduzidas taxas de duplicação. Consequentemente, podemos esperar grandes correlações entre os valores de *ratings* de um programa e a variável dependente.

A segunda abordagem consiste na comparação dos valores de *ratings* de programas adjacentes, sob vários contextos. Quando as correlações são elevadas os autores acabam por concluir que as duplicações de audiências, também, o são (*Davis & Walker, 1990; Tiedge & Ksobiech, 1986*). Apesar das flutuações anuais no decorrer dos anos não se observam grandes alterações nos valores de correlação, adoptando esta tipologia de dados. Este segundo grupo de trabalhos baseia-se em dados de *shares* e de *ratings* divulgados na imprensa da área televisiva. Consequentemente, estes trabalhos identificam-se como estudando apenas as estratégias de *lead_in* ou de manutenção de *shares*, uma vez que os padrões subjacentes de efeitos de herança são essencialmente invisíveis.

O terceiro grupo de trabalhos, assim como o primeiro, medem directamente a audiência partilhada entre programas adjacentes. Em vez de utilizar a percentagem da audiência total que assiste a ambos os programas (*rst*) como variável dependente, utilizam uma medida relativa. Os efeitos de herança são quantificados como a percentagem da audiência de um programa que também assiste a um segundo programa. *Goodhardt et al (1975)* foram os primeiros que descreveram os efeitos de herança desta forma (*rst/rs*). O centro *Nielsen* designa este número por “duplicação primária” das audiências.

Eastman, Newton, Riggs e Neal-Lunsford (1997) utilizam a duplicação primária para avaliar a repercussão de alguns factores na dimensão dos efeitos de herança. Estes autores concluíram que as sinergias tipológicas e o número de programas em contra-programação são responsáveis por 56% da variância explicada e que os efeitos de herança pareciam na altura estar em declínio. No entanto, comparam o modelo que considera como variável dependente a duplicação primária com os modelos que consideram a duplicação total, o que leva inevitavelmente a resultados divergentes (*Webster, 2006*).

O trabalho mais recente que foi possível identificar corresponde ao de *Webster (2006)* que constitui uma réplica das análises realizadas por *Webster (1985)*, utilizando dados de 2004 de duplicações primárias como variável dependente.

Teve como objectivo promover uma melhor compreensão de como as alterações verificadas no meio televisivo norte americano influenciam os efeitos de herança e os factores que explicam este comportamento da audiência. Testa essencialmente quatro hipóteses: (1) Se o *rating* de *lead_in* está positivamente correlacionado com a duplicação, (2) Se o *rating* de *lead_out* está positivamente correlacionado com a duplicação, (3) Se o número de opções está inversamente correlacionado com a duplicação, (4) Se a semelhança entre as tipologias de programas está positivamente correlacionada com a duplicação das audiências. Estas quatro hipóteses são testadas recorrendo a dados de duplicação primária e aos dados de duplicação total (rst) para tentar resolver a discrepância existente na literatura da especialidade a este nível. O autor, também, explora a questão de se os estudos baseados na correlação entre *ratings* adjacentes são na realidade eficazes para explicar os mecanismos de duplicação de audiências.

Do ponto de vista teórico inúmeros autores defendem que o conteúdo programático constitui uma das principais variáveis que determinam o consumo televisivo. Logo, deveria existir algum tipo de lealdade programática ou tipológica, segundo a qual a coincidência ao nível dos conteúdos poderia favorecer a duplicação de audiências. Ao longo das últimas décadas, vários trabalhos de caracterização deste comportamento têm sido desenvolvidos, procurando-se avaliar a influência das sinergias tipológicas entre programas adjacentes nos efeitos de herança, mas não só.

Segundo alguns autores, se o público preferir programas de determinada tipologia é provável que assistam ao programa adjacente se a este estiver associado a mesma tipologia (Davis & Walker, 1990; Headen, Klompmaker & Rust, 1979; Tiedge & Ksobiech, 1986; Webster, 1985). A este nível existem algumas dificuldades na sua avaliação. A primeira e mais relevante será a tipologia programática mais apropriada, existindo consenso de que as categorias devem ser definidas em função das preferências e evitar a utilização das tradicionais. Contudo, nos trabalhos existentes os autores têm recorrido a análises factoriais para promover a sua identificação, mas com pouco sucesso. Segundo alguns autores, se essas categorias existirem, parecem corresponder às tradicionais. Em segundo lugar existe a questão da influência do efeito de saturação (Lehman, 1971).

Este princípio defende a existência de um limite máximo de programas de tipologias coincidentes (ponto de saturação) após o qual o público preferirá seleccionar programas de tipologias complementares. Segundo *Goodhardt et al.* não existe no público uma tendência para seleccionar programas de tipologias semelhantes, uma vez que os padrões de consumo devem-se essencialmente às características de planeamento.

Outro factor a considerar é o número de opções disponíveis na interface de dois programas, o qual deverá penalizar os efeitos de herança (*Davis & Walker, 1990; Tiedge & Ksobiech, 1986; Walker, 1988; Webster, 1985*). Segundo *Walker (1988)*, a tipologia do programa também é relevante, por exemplo, as comédias são os programas que maximizam os efeitos de herança face aos programas subsequentes, enquanto que os programas de mistério, suspense e de variedades são os menos eficazes a este nível. A dimensão da audiência total disponível em determinado instante, também, influencia os resultados e a avaliação da sua repercussão é difícil de concretizar. Uma vez que as mesmas pessoas estão disponíveis e veem televisão em períodos de tempo adjacentes, tem sido difícil separar ambos os factores. Quando a análise se restringe a programas adjacentes a variabilidade observada nas taxas médias de audiências para o total da TV é mais reduzida. Quando as análises incluem programas mais distantes torna-se mais problemático, uma vez que existe um acréscimo na variabilidade associada às taxas médias de audiências para o total da TV. *Tiedge e Ksobiech (1986)* para controlar a influência dos factores de disponibilidade nos efeitos de herança utilizaram dados de *shares*, em vez de valores de *ratings*. Os *shares* são menos sensíveis quanto a diferentes níveis de taxas médias de audiências para o total da TV entre programas, consequentemente será menos provável produzirem efeitos de *lead_in* decorrentes exclusivamente da disponibilidade dos telespectadores.

O modelo de selecção programática proposto por *Webster e Wakshlag (1983)* sugere pelo menos três factores: a disponibilidade da audiência, a estrutura dos programas disponíveis e o conteúdo programático. Defendem que os padrões de disponibilidade da audiência, geralmente independentes da estruturação da grelha de programas, produzem padrões de consumo muito correlacionados. Estes padrões são por sua vez responsáveis pela existência de padrões estruturalmente definidos de fluxos de audiências. Algumas análises realizadas à lealdade aos canais e programática confirmam esta realidade.

A explicação mais plausível para os elevados níveis de duplicação que, por vezes, se observam, é o facto do publico estar disponível e, também, ver televisão em períodos de tempo adjacentes. Em relação à estrutura dos programas disponíveis tem sido difícil avaliar a sua influência, uma vez que as grelhas encontram-se estruturadas por forma a não fazer coincidir o início dos programas transmitidos nos diversos canais. A maioria das vezes, após a conclusão de um programa, o público apenas tem uma opção de continuidade, a não ser que opte por ver um programa que se encontra no seu desenrolar. As preferências tipológicas podem favorecer a sua selecção independentemente do seu início e duração. Assim sendo, as sinergias tipológicas podem favorecer os efeitos de herança e a contra-programação directa e não por complementaridade penaliza-los. Esta teoria do comportamento televisivo entre programas é válida num contexto em que existe um reduzido número de opções, mas, hoje em dia, o processo de escolha tende a ser cada vez mais complexo.

Uma vez que existe uma grande diversidade de alternativas, o público tem dificuldade em identificar todas as opções disponíveis e, também, assistimos a um acréscimo no número de programas da mesma tipologia. Consequentemente, levanta-se a questão de conseguir determinar com exactidão o número de opções disponíveis quando o telespectador individual poderá ou não ter acesso à TV Cabo, ao vídeo ou poderá simplesmente não estar a par de todas as opções disponíveis. *Cooper (1993)* apesar de ter conseguido em parte ultrapassar esta dificuldade, não associou um decréscimo nas correlações dos valores entre programas adjacentes a um aumento no número de opções. Isto indicia que o público com um maior número de opções e com facilidade em seleccioná-las muda de canal para canal entre programas ou durante os programas, anulando um efeito identificável.

O estudo realizado por *Walker (1988)* permitiu concluir que com a penetração do controle remoto, vídeo e da TV Cabo as estratégias baseadas nos efeitos de herança terão cada vez menos sucesso, uma vez que o público será cada vez mais activo. Defende que quando um canal mantém os seus níveis de audiências mais ou menos constantes ao longo de um horizonte temporal, erroneamente se conclui que existem elevados efeitos de herança entre os programas.

Isto não é necessariamente verdade, uma vez que quando se conclui um programa a maior parte da sua audiência poderá mudar de canal, mas simultaneamente processa-se um fluxo inverso vindo dos restantes canais. Em média um programa pode herdar metade da audiência do programa anterior (*Webster, 1985*). Os efeitos de herança são uma das três principais áreas de investigação em que se pode subdividir os trabalhos desenvolvidos no âmbito da duplicação das audiências. As restantes correspondem ao “Repeat Viewing” que se traduz por uma duplicação das audiências dia após dia para os vários episódios de um mesmo programa e na lealdade a programas ou a canais. Em geral, os efeitos de herança ainda permanecem como um fenómeno pouco compreendido. O porquê da sua existência e a influência do conteúdo programático e de outros factores de planeamento a este nível são questões que persistem na actualidade. Consequentemente, recorrendo a dados agregados de audimetria e na ausência de qualquer trabalho que caracterize estes comportamentos a nível nacional caracteriza-se a evolução deste indicador e promove-se a identificação dos factores estruturais que o condiciona, assim como, se avalia a sua relevância para a previsão dos *ratings*.

F.4.2. Objectivos

Neste apêndice pretende-se a análise dos índices de duplicação das audiências entre programas directamente adjacentes, ou não, entre períodos horários e de fluxos de audiências entre canais, considerando várias faixas horárias e diversos factores estruturais. Apenas foi possível identificar trabalhos internacionais que visam a análise de duplicação de audiências entre programas adjacentes no período do horário nobre. Como foi referido por *Webster (2006)*, a maioria dos trabalhos que visam a caracterização deste comportamento recorrem a correlações entre os valores de *ratings* adjacentes e na realidade estas análises não se adequam aos objectivos propostos. Existe uma grande discrepância quanto à forma de medir e analisar os efeitos de herança entre programas adjacentes, levando a resultados contraditórios e muito divergentes mesmo quando visam a análise de dados que incidem no mesmo âmbito geográfico. Do ponto de vista metodológico limitam-se em geral a recorrer a análises de RLM para avaliar a repercussão de um reduzido número de variáveis explicativas a este nível.

Este trabalho pretende a análise de índices de duplicação medidos através das análises individuais do *software* Telereport. Na literatura da especialidade corresponde aos denominados índices de duplicação primária. Deve-se ter em atenção que os dados de duplicação de audiências podem não ser exclusivamente função da passividade da audiência, mas, sim, também influenciados por comportamentos de lealdade ao canal, a tipologias e a conteúdos programáticos. No entanto, é da influência conjunta de todos estes factores que resulta a duplicação das audiências entre programas e diferenciar cada um destes comportamentos será extremamente difícil de concretizar.

Do ponto de vista metodológico apresentam-se perspectivas de análise associadas ao *Data Mining*, algumas das quais se encontram implementadas. Neste capítulo as análises a realizar visam essencialmente os seguintes objectivos:

(1) Avaliação do grau de associação existente entre os valores de *ratings* programáticos globais e parciais de programas adjacentes para cada um dos canais generalistas, procurando diferenciar determinados programas neste âmbito. Pretende-se identificar quais são os programas que mais beneficiam para o período em análise, entre outros factores, das estratégias que pretendem favorecer a duplicação de audiências, como sejam, o “Lead in” e o “Lead out”. Comparam-se os diferentes canais quanto à propensão dos seus públicos para desenvolverem este tipo de comportamento. Com base no grau de correlação existente entre as observações adjacentes avalia-se a viabilidade de recorrer a modelos de Regressão Local Linear para promover a previsão de *ratings*, considerando valores programáticos e associados a faixas horárias.

(2) Caracterização da evolução e variabilidade observada nos índices de duplicação primários entre programas a nível global e parcial; dos índices de duplicação primários entre períodos horários, de fluxos de audiências entre canais e das decisões de ligar e desligar o televisor. Pretende-se avaliar o grau de variabilidade associado destes indicadores para os diversos canais e caracterizar a sua evolução do ponto de vista linear e não linear.

(3) Avaliação da influência de determinados factores de planeamento nos índices de duplicação primária e nos fluxos de audiências para os canais concorrentes, como, por exemplo: o canal de transmissão, as tipologias programáticas envolvidas, os períodos horários e a contra-programação. Pretende-se verificar se a previsão dos valores de *ratings* deve considerar a existência de independência entre os públicos de cada canal e em que medida os factores de planeamento influenciam a ordem de grandeza e a evolução da duplicação de audiências. Será relevante verificar se a duplicação das audiências resulta da lealdade que é dirigida a tipologias e aos canais, ou talvez apenas da lealdade programática e da inércia do público entre programas adjacentes. Pretende-se, também, verificar se o grau de correlação entre valores de *ratings* traduz a real duplicação primária de audiências.

A análise realizada aos índices de lealdade programática demonstrou que apesar de alguns programas apresentarem uma variabilidade não muito significativa dos *ratings* (pelo menos em termos absolutos), esta evolução não pode ser atribuída à existência de elevados índices de duplicação entre episódios de um mesmo programa, ou pelo, menos temporalmente próximos. Os índices em geral não excedem os 50% o que poderá ser atribuído a alterações nos padrões de disponibilidade do painel agravadas pelas frequentes oscilações horárias programáticas. A popularidade de alguns programas não consegue reter os telespectadores quando os programas são transmitidos em períodos mais tardios.

Com base na análise realizada aos valores de *ratings* programáticos, também verificámos que aparentemente existe uma certa independência entre os públicos de cada um dos canais, uma vez que grande parte dos programas dos períodos de maiores audiências apresentam maiores valores de variabilidade como consequência das alterações horárias a que estão sujeitos e apenas da respectiva contra-programação quando está associada a eventos que apresentam uma maior popularidade. Atenuando estas influências os valores de variabilidade, por vezes, decrescem consideravelmente, apresentando valores aceitáveis atendendo às múltiplas fontes nem sempre quantificáveis que podem incutir uma forte variabilidade nestes indicadores.

A variabilidade residual poderá ser atribuída a alterações nos padrões de disponibilidade e na predisposição para ver TV, assim como a fluxos de audiências entre os canais. O painel em função das suas condicionantes profissionais e pessoais poderá não ter disponibilidade para ver TV ou optar por outras actividades de lazer no lar ou no exterior. Não será possível prever estas situações. Atendendo às análises realizadas, os fluxos para os canais concorrentes parecem não ser muito significativos, no sentido de realmente influenciar a evolução dos indicadores de audimetria. O *zapping* pode estar associado a telespectadores muito instáveis que mudam de canal com muita frequência e que veem muito pouca televisão ou a segmentos mais moderados que tendem a mudar de canal com uma menor frequência. Logo, em função do perfil dos *zappers* que possa imperar a nível nacional nos períodos de maiores audiências, a sua repercussão poderá ser ou não significativa na evolução dos *ratings*.

Os canais, também, apresentam reduzidos horizontes temporais no decorrer dos quais conseguem estabilizar os seus níveis de audiências, indiciando a existência de índices de duplicação moderados, uma vez que os *ratings* também resultam da audiência herdada dos canais concorrentes. Este capítulo pretende esclarecer as tendências observadas anteriormente nos dados. Se existirem reduzidos índices de duplicação de audiências em princípio devem-se essencialmente a alterações nos padrões de ver TV e em menor grau a fluxos de audiências mais irregulares entre os canais.

Para o período em análise no início da noite existem maiores valores de *ratings* para a primeira telenovela da SIC e, posteriormente, existe um forte decréscimo nos seus valores a favor do canal TVI. Logo, nestes períodos poderá existir fluxos de audiências entre os canais privados generalistas. Também, foi possível constatar que para a totalidade dos canais existem segmentos horários muito restritos onde possa existir uma maior estabilidade nos níveis de audiências. Estes períodos correspondem essencialmente no horário nobre e ao início da noite e são mais longos para o canal estatal. Os canais privados apresentam uma maior irregularidade na dimensão dos *ratings* para os períodos de maiores audiências, assim como maiores oscilações diárias.

Na RTP1, a selecção é mais selectiva correspondendo quase sempre aos espaços informativos e aos programas directamente adjacentes. Devem existir consideráveis índices de duplicação nestes períodos uma vez que alguns programas apresentam elevados valores de fidelidade. No entanto, deverá ser menos beneficiado por parte de fluxos com origem nos canais concorrentes. Espera-se, também, a existência de consideráveis índices de duplicação para os canais privados e para os períodos de maiores audiências, uma vez que correspondem a períodos onde a audiência pode apresentar uma maior disponibilidade para assistir à TV por um horizonte temporal mais alargado. Deverá existir maiores fluxos de audiências entre os dois canais privados. Quanto à repercussão das sinergias entre tipologias, em geral as grelhas nacionais, nem sempre são estruturadas por forma a favorecer a coincidência tipológica dos programas transmitidos entre períodos adjacentes. Verifica-se essencialmente nos canais privados e no decorrer da noite, transmitindo-se sequências de telenovelas ou séries de ficção. As tardes do fim-de-semana dos canais privados, por vezes, também apresentam sequências de filmes. Atendendo às análises já realizadas em princípio esta variável poderá não apresentar uma grande influência.

Por vezes, no decorrer destes períodos existe um forte decréscimo nos valores dos *ratings* o que indicia reduzidos índices de duplicação. O conteúdo programático e a lealdade programática deverá ter uma maior repercussão nos mesmos. Em princípio a popularidade de determinado programa será um dos principais factores que poderá incentivar à permanência das audiências enquanto aguardam pelo seu início. O canal estatal RTP1 adopta em menor grau por transmitir sequências tipológicas na maior parte da sua emissão. Os canais devem influenciar estes índices, em função do seu grau de notoriedade. Em princípio espera-se que os canais privados maximizem estes valores por um maior horizonte temporal, ainda que a maior irregularidade a que, por vezes, assistimos nos valores dos *ratings* no decorrer dos períodos de maiores audiências indiquem a existência de fluxos entre os canais ou uma maior irregularidade nas decisões de ligar e desligar o televisor. Os períodos horários obviamente que condicionam esta tendência que deverá ser superior em períodos de lazer e função dos dias a que correspondem (semana vs. fim-de-semana). A contra-programação em princípio terá uma maior influência quando se transmitem de programas de maior popularidade.

Em *Fernandes (2001)* encontramos observações que levantam inúmeras incógnitas e que se pretendem avaliar nas análises a realizar: “... As mulheres, sobretudo as que não trabalham, são as pessoas que veem mais televisão durante o dia. Por outro lado, as audiências de televisão variam em diferentes períodos do dia, mas coincide em períodos adjacentes.” Será necessário verificar se nas faixas horárias anteriores ao pré-prime-time a duplicação de audiências é realmente superior. Uma das questões a desenvolver corresponde à influência dos horizontes temporais na duplicação das audiências.

A mesma autora refere que: “ ... a duplicação da audiência pode afectar a audiência de qualquer outro programa – do primeiro programa *lead_in* uma boa percentagem de audiência vai para o segundo.” Será necessário constatar se realmente a duplicação de audiências entre programas adjacentes é considerável. Na bibliografia da especialidade apenas se avalia este indicador desde um programa para o subsequente. Segundo *Fernandes*, o *zapping* varia segundo o tempo de visionamento da televisão, ao longo da semana e, em particular a partir das 20.30H. Salienta algumas propostas de outros autores, entre as quais a existência de uma tipologia de telespectadores que muda de canal várias vezes: os instáveis que mudam frequentemente (entre 3 a 5 minutos) e veem pouca televisão; os infiéis com o mesmo perfil ainda que as mudanças sejam menos frequentes; os moderados; os desinteressados que veem um pouco mais de TV e os contemplativos que correspondem ao segmento que muda menos de canal e que abrange essencialmente as pessoas inactivas e idosas.

Com base nas análises a realizar no *Apêndice F.5* será possível avaliar o perfil sociodemográfico de cada canal em função de diferentes faixas horárias e para os dias laborais versus fim-de-semana, verificando se existem divergências face ao que consta em *Fernandes (2001)*. Segundo *Fernandes*, o público da RTP é rural e idoso, logo a duplicação de audiências deveria ser superior para este canal face aos canais privados no decorrer da totalidade da emissão diária, o que se irá avaliar nas análises a realizar (corresponde ao segmento que é designado de contemplativos). A autora salienta que o canal SIC no decorrer do dia apresenta uma audiência de donas de casa e a partir do horário nobre todos os segmentos estão presentes.

Logo, a duplicação de audiências, também, deverá ser razoável no decorrer do dia e diferenciar-se em função das diversas faixas horárias e, por último a TVI é associada por *Fernandes (2001)* aos jovens dos 15 aos 40 e às classes sociais mais altas, logo a duplicação de audiências deverá ser maximizada neste canal. Estas são algumas das questões a constatar nas análises a realizar.

F.4.3. Bases de dados

A base de dados inclui programas repartidos ao longo de toda a semana e períodos horários e dos quatro principais canais, ou seja, a SIC, TVI, RTP1 e Canal 2. Estes índices têm sido medidos de três formas distintas e que constam do *Quadro F.4.3.1*.

Taxas de Audiências Condicionadas
A percentagem da audiência de um programa que, também, assiste ao programa adjacente. Podemos calcular estes valores através das análises individuais do <i>software</i> Telereport. Esta tipologia de dados representa valores relativos e como tal são insensíveis aos valores de <i>ratings</i> dos programas envolvidos em cada um dos pares considerados nas análises.
Taxas de Audiências Conjuntas
A percentagem da audiência total que assiste a ambos os programas. Não é possível extrair este valor directamente através dos <i>softwares</i> disponíveis, mas pode ser calculado multiplicando os valores da duplicação primária e os valores de <i>ratings</i> . Estes valores podem ser interpretados como medidas absolutas da dimensão da audiência e podem ser calculados com os dados que constam das bases de dados de audimetria.
Valores de Ratings
De forma indirecta comparando o <i>rating</i> de um programa com o <i>rating</i> do programa anterior ou sucessor. Esta última hipótese apresenta a vantagem de se poder medir este efeito quando apenas estão disponíveis dados agregados de valores de <i>ratings</i> ou valores de <i>shares</i> , mas em <i>Webster (2006)</i> demonstra-se que esta perspectiva de análise não é certamente a mais correcta.

Quadro F.4.3.1 - Perspectivas de quantificação da duplicação das audiências que permitem avaliar os efeitos de herança entre os suportes publicitários

Alguns autores defendem que nenhuma das tipologias de dados anteriormente enumeradas pode ser encarada como sendo a mais adequada às análises propostas. Porém, para *Webster (2006)*, a primeira opção é a mais adequada e será a variável critério a adoptar na maioria das análises a realizar nesta tese.

Para a diferenciação dos programas quanto aos que mais beneficiam da popularidade dos programas adjacentes irá medir-se os índices de duplicação de forma indirecta, utilizando o *rating* do programa anterior (*Rat_Ldin*) e o *rating* do programa posterior (*Rat_Ldout*) (dados referentes à 3ª Opção), uma vez que os dados relativos são insensíveis à popularidade dos programas envolvidos nas análises. Um elevado índice de duplicação não corresponderá necessariamente a programas que por si só estejam mais expostos ao público. Para a caracterização e posterior avaliação da repercussão dos factores de planeamento nos índices de duplicação recorre-se aos quadros de índices de duplicação primários (opção 1) onde uma célula (i,j) representa a proporção da audiência que assiste ao programa em linha (i) e que também assiste ao programa em coluna (j). Estes quadros podem caracterizar os programas globais (nível 1 programático), ou também se poderá considerar cada uma das suas partes (nível 2) (ver, por exemplo, *Quadro F.4.3.2*).

Dia	Hora	Canal	Nome Prog j ₁	...	Nome prog j _n
Nome Prog i ₁			P(j ₁ /i ₁)		P(j _n /i ₁)
....					
Nome Prog i _n			P(j ₁ /i _n)		P(j _n /i _n)

Quadro F.4.3.2 – Índices de duplicação de audiências (*software Telereport*)

Dia	Id. Prog. de Referência	Duplicação de Audiências
Dia I	Nome Prog i1	Nome Prog j1 P(Prog j1/ Prog i1)
	Nome Prog i1	Nome Prog j2 P(Prog j2/ Prog i1)
	...	...
	Nome Prog i1	Nome Prog jm P(Prog jm/ Prog i1)
	...	...
	Nome Prog in	Nome Prog j1 P(Prog j1/ Prog in)
	Nome Prog in	Nome Prog j2 P(Prog j2/ Prog in)
	...	...
	Nome Prog in	Nome Prog jm P(Prog jm/ Prog in)

Quadro F.4.3.3 – Base de dados de índices de duplicação das audiências

Através das análises individuais obtemos um quadro para cada dia seleccionado, onde apenas consta a identificação do programa, dia e hora de transmissão.

Para construir uma base de dados apta a ser analisada será necessário agregar os dados de cada um dos dias (*Quadro F.4.3.3*) e tomar como unidade de análise cada par de programas que consta dos referidos quadros.

As bases de dados das análises individuais do *software* Telereport não estão inter-relacionadas com os índices de audiências programáticos, não sendo viável por meio do *software* Telereport agregar a estas bases outras variáveis de caracterização da programação. Posteriormente, essas variáveis foram adicionadas, através da base de dados construída para o efeito. Consideram-se índices parciais e globais de cada um dos programas. A partir destes valores, também se irá construir uma nova base de dados na qual os índices se referem a períodos horários com uma duração de 15' em substituição dos programas em concreto (Quadro F.4.3.4).

Dia	Id. Supurtes de Referência	Restantes Suportes	Duplicação de Audiências
Dia I	8-8.1Q	8-8.2Q	X_1
....		8-8.3Q	
....		8-8.4Q	
....			
....			
....		24-24.1Q	
....		24-24.2Q	
....		24-24.3Q	
....		24-24.4Q	X_m

Os dados para períodos horários, foram calculados com base nos índices de duplicação primários parciais associados aos programas que são transmitidos nesses períodos. Assim como se verificou nos capítulos anteriores as análises referentes a programas apresentam consideráveis acréscimos na variabilidade dos resultados como consequência das oscilações horárias observadas.

Quadro F.4.3.4 – Base de dados de índices de duplicação para as faixas horárias

Na literatura da especialidade não foi possível identificar trabalhos que recorram aos dados de períodos horários. A unidade estatística corresponde a pares de períodos com 15 minutos de duração, por exemplo, 19H–1Q / 19H–2Q, que representa o primeiro e o segundo quarto de hora das 19 horas. Os índices de duplicação primários não têm em atenção, por si só, o horizonte temporal que cada telespectador permanece a eles sintonizado. Traduzem a percentagem da audiência que viu determinado programa independentemente do tempo de contacto e foram seleccionadas duas opções para a definição de telespectador. Numa base de dados considerou-se um segundo de modo a incluir a totalidade da audiência e numa segunda base de dados como sendo uma exposição superior a 50% da duração do programa ou período considerado.

Será necessário uma significativa prévia reestruturação das bases de dados, uma vez que os dados constam de quadros diários que são independentes e não inter-relacionados onde se cruzam os vários programas e onde, também, não constam as variáveis de caracterização dos programas envolvidos nas análises nem da sua envolvente (adjacentes e concorrentes).

A estas bases de dados podem-se, também, agregar os dados de fluxos de audiência entre os canais, os quais fazem parte de outra tipologia de análises individuais do *software* Telereport. Para cada um dos programas seleccionados existem índices de duplicação face aos restantes programas da grelha e índices que quantificam os fluxos que se verificam desde esse programa para os canais concorrentes ao longo de toda a grelha de programação, assim como a percentagem da audiência que permanece com o televisor ligado ou que o desliga.

Correspondem uma vez mais a dados diários que devem ser agregados considerando como unidade estatística cada par de programas. Não é viável, através do *software* Telereport, adicionar algumas variáveis de caracterização da programação envolvida nestas bases de dados (*Quadro F.4.3.5*).

	SIC	TVI	RTP1	RTP2	TV_Off	TV_On	TV_Cabo
Prog 1	P (SIC/ Prog 1- RTP)	P (TVI / Prog 1- RTP)	100%	P (RTP2 / Prog 1- RTP)	P(Tv_Off/ Prog 1- RTP)	P(Tv_On/ Prog 1- RTP)	P(Tv_Cabo/ Prog 1- RTP)
Prog 2	P(SIC (no decorrer de prog 2) / Prog 1 – RTP1)	P(TVI (no decorrer de prog 2) / Prog 1 – RTP1)	P (Prog2 / Prog1)	P(RTP2 (no decorrer de prog 2) / Prog 1 – RTP1)	P(TV_Off (no decorrer de prog 2) / Prog 1 – RTP1)	P(TV_On (no decorrer de prog 2) / Prog 1 – RTP1)	P(TV_Cabo (no decorrer de prog 2) / Prog 1 – RTP1)
...	...	...	...	...	...	...	...
ProgN	P (SIC/ Prog 1- RTP)	P (TVI/ Prog 1- RTP)	P (RTP1 / Prog 1- RTP)	P (RTP2/ Prog 1- RTP)	P (TV_Off/ Prog 1- RTP)	P (Tv_On/ Prog 1- RTP)	P (Tv_Cabo/ Prog 1- RTP)

Quadro F.4.3.5 – Índices de fluxos de audiências entre suportes horários ou programáticos

A partir destes quadros diários, também, é necessário construir uma segunda base de dados tomando como referência pares de programas (*Quadro F.4.3.6*).

O que se pretende não se pode traduzir apenas numa análise exclusiva dos efeitos de herança, mas, sim, na caracterização dos índices de duplicação entre programas independentemente da sua motivação, origem e duração.

Índices de Fluxos			
Dia I	Prog i	Progi-n	...
...	...	Progi-1	...
...	...	...	...
...	...	Progi+1	...
...	...	...	...
...	...	Prog_N	...
...	...	...	...

Quadro F.4.3.6 – Base de dados de fluxos de audiências

A nível prático será o valor global que poderá condicionar o desempenho da grelha de programas. Estas análises são relevantes na medida em que permitem a identificação de factores que podem levar os canais a maximizar e estabilizar os seus níveis de audiências por um maior horizonte temporal, facilitando a previsão dos valores de *ratings*.

Relativamente às variáveis explicativas vamos considerar variáveis de caracterização da grelha de programação e determinados índices de audiências.

(1) Variáveis de caracterização da programação: Para cada par de programas consideramos a sua hora, duração global e parcial, número de partes que integram cada programa, tipologias programáticas de grau 1, 2 e 3, sinergias tipológicas, durações relativas, dia de semana, entre outras, inúmeras variáveis que tiveram que ser calculadas.

(2) Índices programáticos de audiências globais e parciais: Consideramos a totalidade dos índices programáticos de ambos os programas incluídos em cada par de observações, por exemplo, os respectivos valores de *ratings*, *shares*, taxas médias de audiências para o total da TV, índices de saturação publicitária, entre outros. Consideram-se, também, para alguns indicadores os valores relativos que resultam do quociente entre os valores associados a ambos os programas.

F.4.4. Metodologias

I. Caracterização dos índices de duplicação associados a programas

Para além das análises gráficas que procuram identificar tendências, uma vez que um dos objectivos das análises reside na caracterização da variabilidade dos índices de duplicação, é viável recorrer à Análise de Correspondências e/ou a análises MDS para tentar identificar quais são os programas que apresentam maiores duplicações de audiências médias para o horizonte temporal em análise.

As unidades estatísticas correspondem a pares de programas adjacentes, mas, não só, caracterizados pelos valores médios dos índices de duplicação primários. Consideram-se apenas os valores para o universo televisivo, mas diferenciam-se os resultados para cada um dos canais.

II. Caracterização de índices de duplicação primários associados a faixas horárias

Atendendo aos dados de duplicação primária para períodos horários e considerando os valores médios observados, também podemos recorrer às duas metodologias enumeradas anteriormente, uma vez que os índices representam as probabilidades condicionadas de selecção de dois períodos horários $P(P(t_{i+1}|t_i))$. As probabilidades condicionadas podem representar, também, o grau de semelhança existente entre o público de dois períodos horários.

A análise de Correspondências e a análise MDS permitem-nos a visualização dos períodos que perdem um maior número de telespectadores na sua interface por decisões de desligar o televisor ou como consequência da existência de fluxos de audiências para os canais concorrentes. As unidades estatísticas correspondem a pares de intervalos de 15 minutos caracterizadas pelos correspondentes valores médios dos índices de duplicação primária.

Nas análises à variabilidade destes indicadores recorre-se a análises gráficas. Segundo os resultados e a disponibilidade temporal avaliar-se-á a necessidade e a viabilidade de recorrer a metodologias equivalentes simbólicas que têm em consideração os intervalos de variabilidade inerentes às variáveis em estudo. Os períodos de referência correspondem às faixas horárias após as 19H e à hora do almoço.

III. Avaliação da repercussão dos factores de planeamento na duplicação das audiências

Após a caracterização da evolução e variabilidade observada nestes indicadores procede-se à avaliação da repercussão de determinados factores de planeamento. Os trabalhos publicados recorrem essencialmente a análises de Regressão Linear Múltipla que não são as mais adequadas, uma vez que apenas avaliam a existência de relações lineares entre as variáveis em estudo. Uma vez que as relações podem não ser lineares parece ser mais relevante recorrer a metodologias de *Data Mining* não lineares, como, por exemplo, às Árvores de Regressão que permitem a identificação das interacções mais significativas entre as variáveis.

Após a aplicação das Árvores de Regressão em função dos resultados avaliar-se-á a necessidade de recorrer a metodologias da área do *Temporal Reasoning*. Algumas sugestões metodológicas são apresentadas para a análise dos índices de duplicação. A representação dos dados de audimetria em séries temporais simbólicas, também, viabiliza a implementação de metodologias associadas ao *Temporal Reasoning*.

IV. Quantificação da Duplicação das Audiências entre suportes publicitários

Numa tentativa de quantificação da duplicação de audiências entre programas podemos recorrer à Teoria da Probabilidade, e em especial à Estatística Bayesiana como consta do *Capítulo VI.4.4.1*.

F.4.5. Resultados

Os resultados apresentam-se subdivididos em cinco partes.

(1) A primeira parte avalia indirectamente os efeitos de herança, recorrendo aos dados de *ratings*. Verificamos qual é o grau de correlação linear existente entre as observações adjacentes e avalia-se em que medida a correlação traduz o grau de duplicação primária de audiências.

(2) Na segunda parte consta a caracterização da evolução dos índices de duplicação primários para os três principais canais generalistas.

(3) Na terceira parte associam-se os índices de duplicação a períodos horários e avalia-se o seu grau de variabilidade, recorrendo-se à Análise de Correspondências e à Análise MDS. Agregou-se às bases de dados algumas variáveis que caracterizam os fluxos de audiências entre os canais e as decisões de ligar e desligar o televisor.

(4) Na quarta parte avalia-se a repercussão dos factores estruturais nos índices de duplicação primários dos períodos horários e dos programas que podem ser adjacentes ou não. Correspondem a análises de Regressão Linear Múltipla e a Árvores de Regressão não Linear. As análises podem se referir à totalidade da emissão diária ou apenas aos períodos de maiores audiências.

(5) Por último tenta-se promover a quantificação dos índices de duplicação entre programas recorrendo à teoria da probabilidade e que também consta do *Capítulo VI.4.4.1*.

Também, se salientam algumas abordagens metodológicas que podem contribuir para a previsão de valores de *ratings*, considerando a exposição aos canais e à duplicação de audiências.

I. Análises de Duplicação de Audiências Recorrendo aos *Ratings* Programáticos

Pretende-se avaliar a seguinte questão: Qual o grau de correlação entre os valores de *ratings* de programas adjacentes? Qual a viabilidade de recorrer a Regressões Locais para prever os valores de *ratings* programáticos ou para faixas horárias? Em que medida os canais influenciam os resultados?

Considerando a globalidade dos dados, o *rating* de um programa apresenta uma correlação ligeiramente superior com o *rating* do programa sucessor (*Rat_Ldout%*, $\rho=0,68$ e *Rat_Ldin%*, $\rho=0,62$). Aparentemente existe uma maior associação entre o desempenho de um programa e o do seu programa sucessor, pelo menos do ponto de vista linear. Contudo, o diferencial de valores é pouco significativo. Se em substituição considerarmos os valores de *ratings* dos períodos adjacentes para cada um dos canais, os coeficientes de correlação são de 0,88 para *Rat_Período_Ldout%* e de 0,88 para *Rat_Período_Ldin%* considerando os dados do canal SIC; 0,81 e 0,82 respectivamente para a RTP1 e de 0,93 e 0,95 para a TVI desde que não excluindo os programas que geram observações atípicas. Estes valores são muito superiores aos anteriores, não se diferenciando os valores de *ldin* face aos valores de *ldout*.

Excluindo programas de maior popularidade os valores são respectivamente de 0,88 e 0,89 para a SIC, 0,85 e 0,86 para a RTP1 e de 0,93 e 0,95 para a TVI. Logo, existe linearidade e consequentemente proporcionalidade entre os valores de *ratings* de observações adjacentes se considerarmos os valores de períodos horários, não se diferenciando significativamente os valores de *Ldin* face aos valores de *Ldout*. Provavelmente estes valores podem-se diferenciar se considerarmos períodos horários mais restritos. Nas análises considerou-se intervalos de meia hora.

No entanto, as análises de Regressão local Linear realizadas com valores de *Tvr_Período%* apresentam no máximo 75% de variância explicada. Com base nestes valores de correlação seria de esperar valores mais elevados e as metodologias de Regressão local linear ou localmente polinomiais deveriam constituir uma boa opção para tentar prever os valores de *ratings* de faixas horarias, sendo posteriormente necessário associar os valores previstos aos programas.

Considerando apenas os dados do canal SIC e em relação aos valores de *ratings* dos programas adjacentes a variável *Rat_Ldout%* supera *Rat_Ldin%* (*Rat_Ldout%*, $\rho=0,62$ e *Rat_Ldin%*, $\rho=0,52$). Logo, existe uma maior proximidade de valores face aos valores de *ratings* dos programas sucessores e a duplicação das audiências, em princípio, também, é mais frequente em sentido directo, favorecendo os programas subsequentes. A variável *Rat_Ldin%* apresenta um menor coeficiente, uma vez que o fluxo de audiências em sentido inverso deverá ser menor, pelo menos do ponto de vista linear.

Uma vez que as Regressões Locais consideram as K observações mais próximas para estimar a recta de Regressão, as observações que antecedem determinado programa contribuem para minimizar o desempenho global dos modelos locais anteriormente calculados.

Para o canal TVI, os *ratings* dos programas adjacentes equiparam-se quanto aos valores dos coeficientes de correlação que apresentam (*Rat_Ldout%*= *Rat_Ldin%*=0,76) o que poderá ser atribuído a oscilações menos bruscas nos índices de audiências entre os programas adjacentes.

Por último, para a RTP1 os coeficientes de correlação são de 0,71 para *Rat_Ldout%* e de 0,63 para *Rat_Ldin%*, coincidindo com a tendência identificada anteriormente para o canal SIC. Aparentemente, a duplicação entre programas adjacentes é inferior para este canal, existindo uma selecção programática mais selectiva e pontual. Os espaços informativos impulsionam os *ratings* dos programas adjacentes.

II. Análises Descritivas aos Índices de Duplicação Primários associados a Programas

Uma segunda forma de medir os efeitos de herança consiste em recorrer às taxas de audiências condicionadas entre dois programas. As seguintes análises baseiam-se nesta tipologia de dados, adoptando-se o primeiro nível programático, ou seja, a sua duração total. Os resultados diferenciam-se para cada um dos canais generalistas.

Posição	1	2	3	4	5	6	7	8	9	10
1	100	64,66	51,26	44,46	41,24	33,80	29,15	29	15,30	8,6
2	58,05	100	69,80	55,36	45,90	35,56	28,05	23,50	10,80	2,3
3	49,81	72,25	100	68,58	53,55	40,31	29,42	19,47	6	1
4	44,15	59,08	68,15	100	69,22	50,54	38,35	22,23	5	5,6
5	33,90	40,79	44,84	57,94	100	62,95	43,46	28,63	12,2	2
6	21,45	24,47	25,78	31,45	45,82	100	63,76	41,83	18	18,6
7	11,46	11,78	11,62	14,40	19,02	34,28	100	59,53	32,9	23,30
8	14,27	13,63	11,60	15,70	17,03	21,20	44	100	46,8	30
9	11,20	12	9,60	8,10	12,90	10,90	19,20	25,80	100	58,20
10	8	4,70	2,2	9	8,50	5,3	10,40	10,60	33,70	100

Quadro F.4.5.II.1 – Média de valores de duplicação de audiências em função da posição dos programas e para o canal SIC

Como foi anteriormente referido, estes dados apresentam-se em forma matricial, cruzando os vários programas que constituem a grelha diária de programação. Através do *software* Telereport obtemos um quadro para cada um dos dias seleccionados. A partir dos quadros diários calcularam-se os valores médios da variável em estudo para os programas envolvidos em cada par considerado ou em função das respectivas posições na grelha de programas independentemente de quais são os programas transmitidos.

No Quadro F.4.5.II.1 constam as taxas de audiência condicionadas médias em função da posição do programa na grelha e apenas para o canal SIC, a título ilustrativo. A primeira posição corresponde ao espaço informativo da noite. Verificamos que o programa sucessor do espaço informativo da noite da SIC herda em média 58% da audiência desse programa (posição 2,1).

Apesar da frequente popularidade da primeira telenovela da noite quando um programa é intercalado entre este programa e o espaço informativo assistimos, por vezes, a uma ligeira redução nos *ratings*, apesar dos índices da telenovela, geralmente, superarem os valores do espaço informativo. Portanto, neste período as pessoas podem optar por ver um canal concorrente ou por realizar outras actividades.

Existe um segmento da audiência que apenas liga o TV no canal SIC para ver o espaço informativo da noite. Em média 42% dos telespectadores desliga o televisor ou muda de canal. Para este segmento a análise dos índices de lealdade programática será mais relevante e uma proposta metodológica de previsão consta do capítulo VI. A duplicação das audiências face aos períodos posteriores, também, vai decrescendo no decorrer da noite de forma muito significativa. No final da noite apenas 11.2% em média dos telespectadores que assistiram ao espaço informativo ainda apresentam o televisor ligado e a assistir a programas da SIC. Esta percentagem é mais elevada (72%) considerando a audiência conjunta dos segundos e terceiros programas da grelha de programas da noite. Nestes horários para alguns dos anos em análise existe uma grande oscilação na ordem e, consequentemente, nos horários dos programas que geralmente são transmitidos.

Mas, em alguns anos existe uma maior estabilidade, por exemplo, em 2012 verifica-se alguma estabilidade horária o que simplifica as análises. Quem opta por ver o programa de ficção sucessor do espaço informativo da noite, também, apresenta uma maior probabilidade de assistir ao programa de ficção subsequente, o que também pode ser associado a uma lealdade tipológica e neste caso às telenovelas.

Por efeito de “lead_in” ou de “lead_out”, sem dúvida que um programa com maiores níveis de exposição favorece parcialmente a passividade da audiência nos períodos que lhe são adjacentes. Portanto, o público que decide aguardar pelo início do programa que pretende sem mudar de canal terá obviamente que apresentar elevados índices de duplicação face ao programa de sua eleição. Este período horário abrange parte do que tradicionalmente se define como horário nobre e, também, o início da noite, e, logo, um horizonte temporal onde grande parte do painel pode apresentar uma maior disponibilidade para assistir à TV durante um maior horizonte temporal.

A percentagem de duplicação de audiências associada aos terceiros e quartos programas da grelha diminui ligeiramente a nível médio (68,15%). A nível horário abrange parte do final da noite, mas a descida face ao período anterior é de apenas 4%. O carácter narrativo das telenovelas e a atractividade que as telenovelas brasileiras e, posteriormente as nacionais, também geram contribui para incentivar à duplicação de audiências.

A partir da quarta e sucessivas posições começamos a observar uma redução muito significativa nestes valores. Obviamente, como consequência da hora uma vez que vamos progredindo pelo final da noite e madrugada. Portanto, os programas que começam após o espaço informativo da noite e que têm início até por volta das 22.30H são os programas que na realidade recebem mais audiência do programa que lhes antecede. Os resultados apresentados resultam de considerar apenas os programas directamente adjacentes.

Se considerarmos a evolução destes indicadores no decorrer da noite, as maiores reduções de programa para programa verificam-se a partir dos programas mais afastados do início do horário nobre.

O público que assiste ao espaço informativo da noite é o que aparentemente mais lentamente muda de canal ou desliga o televisor. Os telespectadores que começam a assistir à programação da SIC em períodos posteriores são, também, os que mais rapidamente o podem deixar de fazer. Provavelmente, também por seleccionarem os canais concorrentes, uma vez que apresentam uma maior lealdade programática ou tipológica, mas uma menor ligação para com o canal. Os telespectadores que assistem ao espaço informativo da noite deste canal são os mais leais e, consequentemente apresentam uma menor predisposição para mudar de canal no decorrer da noite.

III. Análises aos índices de duplicação entre os períodos horários

Nas análises a realizar neste item recorreremos aos dados referentes a períodos horários. As publicações que foi possível identificar referem-se apenas a valores programáticos e adjacentes. Uma vez que os programas apresentam grandes oscilações horárias pretendemos, assim, atenuar a variabilidade observada nestes indicadores. As unidades estatísticas correspondem a pares de períodos de 15 minutos, por exemplo, 19/1Q-19/2Q que representa o primeiro e segundo quarto de hora das 19H. Adicionou-se às bases de dados, as variáveis que traduzem os fluxos de audiências existentes entre os canais e as decisões de ligar ou desligar o televisor.

Recorre-se à Análise de Correspondências, análise de *Clusters* e a análises MDS para identificar quais são os períodos horários que partilham uma maior percentagem de audiência.

IIIa. Análises aos valores médios dos índices de duplicação e de fluxos de audiências entre canais para o horário nobre e noite do Canal SIC

Calcularam-se os valores médios dos índices de duplicação primária desde o primeiro quarto de hora das 19H até à 1H, de 15 em 15 minutos.

Realizando uma análise de *clusters* hierárquica com base na matriz que representa as percentagens de duplicação médias de audiências entre os períodos horários verificamos que sobressaem os seguintes conjuntos de períodos: (a) 19/1Q- 20/3Q, (b) 20/4Q-22/3Q e (c) 0/1Q-0/4Q.

Portanto, as faixas horárias compreendidas entre os períodos 22/4Q e 23/4Q, correspondem a um período de transição em que os índices de duplicação são cada vez menores de período para período.

Atendendo aos valores médios dos índices de duplicação realizou-se uma análise MDS e de Correspondências (*Figura F.4.5.III.1*).

Os índices de duplicação representam ocorrências entre dois períodos e, consequentemente adequam-se a estas metodologias. Os resultados das análises MDS são mais lógicos e correspondem aos que resultam das análises gráficas realizadas.

Com base na análise MDS onde se cruza as dimensões 1 e 2 verificamos que a partir das 23H vai existindo um distanciamento cada vez maior entre os períodos, o que podemos atribuir aos fluxos direccionados para os restantes canais, assim como às decisões de desligar o TV. De quarto em quarto de hora existe um decréscimo muito significativo nos índices de duplicação, uma vez que corresponde à conclusão da telenovela do horário nobre.

Identificamos uma maior proximidade entre os períodos 19H a 20.30H que representa o pré-prime-time e a primeira parte do espaço informativo da noite, 20.30H a 21H que representa o fim do espaço informativo, 21H a 22H (representa as faixas horárias sucessoras ao espaço informativo e que em alguns anos antecede a telenovela do horário nobre. Em 2012, por exemplo, a SIC opta por transmitir uma telenovela logo após a conclusão do espaço informativo), 22.15H a 23H (representa a telenovela do horário nobre ou a segunda da noite em função dos anos). Os maiores valores estão restringidos a intervalos de muito curta duração. Se o canal SIC alterar a grelha de programas de maneira a prolongar um pouco mais o seu período de maiores audiências, obviamente que os índices de duplicação seriam superiores nestas faixas horárias.

A análise de Correspondências não permitiu alcançar bons resultados.

Estas análises, também, podem ser realizadas com os valores programáticos. Os programas que nos mapas MDS se localizam com uma maior proximidade partilham um maior número de telespectadores. Se existir dispersão de períodos temporalmente próximos, obviamente que as reduções nos índices podem resultar de decisões de desligar o televisor ou da existência de fluxos de audiências para os canais concorrentes.

Em princípio espera-se um maior distanciamento entre os períodos a partir das 23H, uma vez que segundo as análises realizadas aos valores de *ratings*, corresponde ao ponto a partir do qual se verifica uma tendência descendente nos seus valores.

Nos gráficos *F.4.5.III.2* a *F.4.5.III.7* está representada, não só, a evolução média dos índices de duplicação, como também dos fluxos de audiências entre os canais. Estes gráficos ilustram a evolução dos índices de duplicação face a determinados períodos horários de referência (por exemplo, 19/1Q - 0/1Q – 1º quarto de hora das 19H até ao 1º quarto de hora das 24H).

□ **Período de Referência: 19H/ 1Q**

Com base nas análises gráficas que traduzem as evoluções médias dos índices de duplicação e dos fluxos de audiências face a determinados períodos de referência verificamos que tomando como referência 19/1Q existe uma evolução linear até às 24H na evolução da variável que traduz a percentagem do público que permanece com o televisor ligado (Tv_On) e representada pela $y = -3,3x + 93,3$, $R^2 = 0,97$.

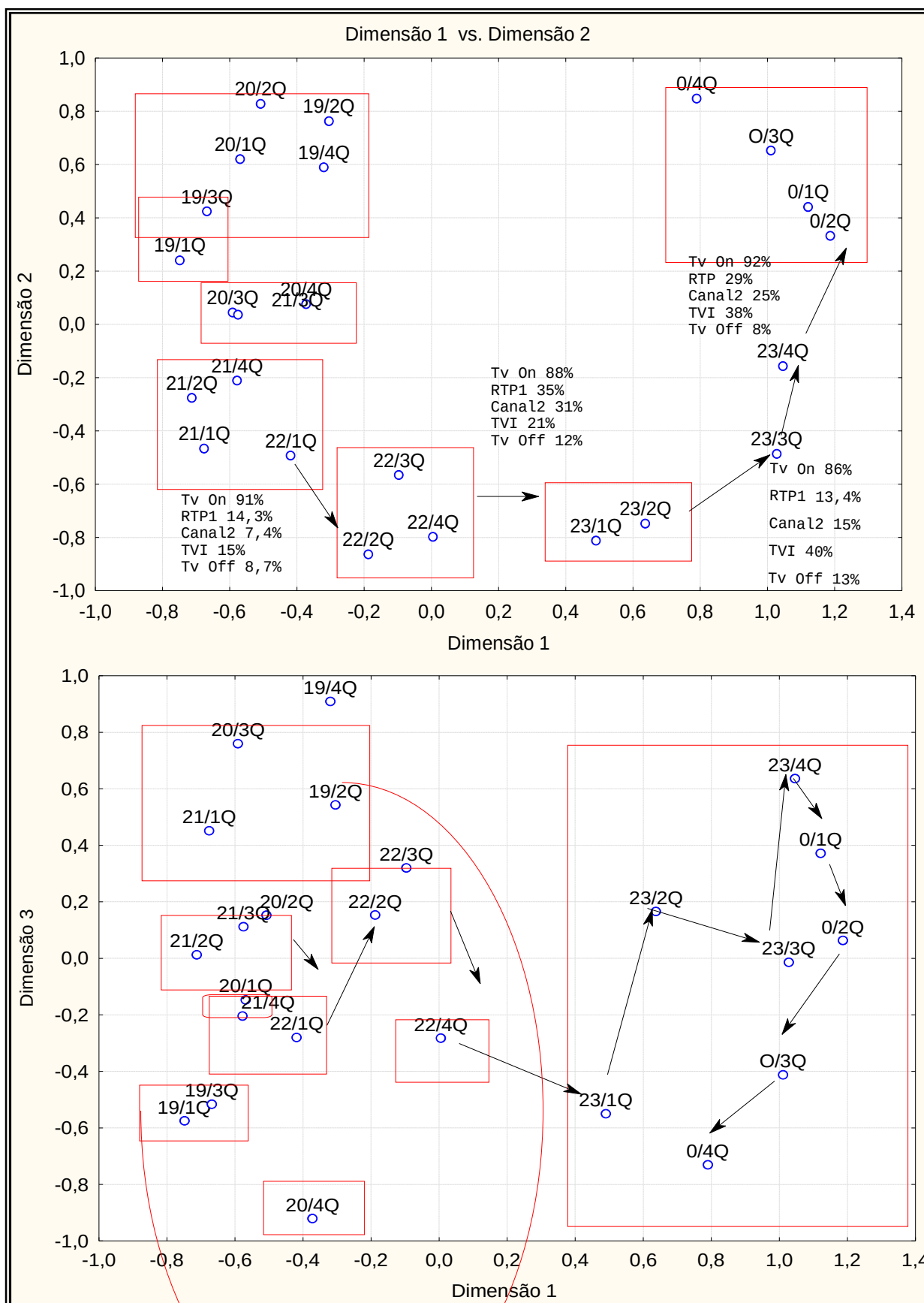


Fig. F.4.5.III.1 – Resultados da Análise MDS para as dimensões 1, 2 e 3

Por volta das 23H verifica-se um cruzamento entre as linhas *Tv_On* e *Tv_Off*, ou seja, corresponde ao início do período em que grande parte do painel opta por desligar o televisor, produzindo-se um decréscimo muito significativo nos valores dos índices de duplicação e, por conseguinte nos valores dos *ratings*. A recta que define a evolução dos índices de duplicação é paralela à recta *Tv_On*, excluindo os períodos 21/1Q-21/2Q onde se verifica um maior decréscimo nos seus valores (+/-10%).

Esta descida poderá ser atribuída a fluxos ou a decisões de desligar o televisor. Neste caso, prevalece a segunda opção, ou seja, quem assiste à TV desde o início do pré-prime-time, por vezes, aproveita o final do espaço informativo da noite para realizar outras actividades. Logo, são as decisões de manter o televisor ligado que condicionam em maior grau a evolução da duplicação de audiências e, por conseguinte, também dos *ratings*.

Em relação aos valores dos fluxos de audiências face aos canais concorrentes a TVI supera ligeiramente a RTP1 em grande parte do período em análise, existindo para ambos os canais um acréscimo mais significativo no decorrer do espaço informativo da noite.

Os valores de fluxos para os restantes canais são irrisórios. Interessa realçar que os valores dos fluxos apenas representam a percentagem da audiência que selecciona pelo menos um segundo os programas dos canais concorrentes.

Logo, não implicam obrigatoriamente uma redução nos índices de duplicação e nos valores dos *ratings*.

□ Período de Referência: 21H/ 1Q

Com base no gráfico que adopta 21/1Q como referência (*Gráfico F.4.5.III.5*) verificamos uma vez mais o paralelismo que existe entre a variável *Dup_SIC* e *Tv_On*. O maior decréscimo nos índices de duplicação corresponde ao ponto em que *Tv_On* e *Tv_Off* se cruzam e que corresponde à primeira meia hora das 23H. Os valores referentes aos fluxos para os restantes canais são inferiores aos 30%. Com base no gráfico que adopta 22/1Q como referência os resultados são semelhantes (*Gráfico F.4.5.III.4*). As rectas *Tv_On* e *Tv_Off* cruzam-se nos mesmos períodos e é notório o decréscimo muito significativo que se verifica nos índices de duplicação após a conclusão do espaço informativo da noite acompanhado de uma idêntica evolução nos valores da variável *TVI* e *RTP1*. A partir das 23H existe uma cada vez maior proximidade entre a recta associada à *SIC* e a *Tv_On*, ou seja, os índices de duplicação não resultam apenas das decisões de desligar o TV, mas, também, dos fluxos para os canais concorrentes. Simultaneamente, verifica-se uma maior proximidade entre as rectas referentes ao canal *SIC* e da *TVI*, ou seja, existe um fluxo muito significativo de telespectadores para a *TVI*. Consequentemente, verificamos que no decorrer da noite a lealdade programática é mais acentuada do que a lealdade dirigida aos canais.

As pessoas seleccionam os programas, independentemente, do canal a que pertencem. Para o público que sintoniza o TV às 19/1Q os fluxos que existem no final da noite entre a TVI e SIC não são acompanhados por uma evolução semelhante para com a RTP1. No público que sintoniza o TV a partir das 21/1Q é mais notória a proximidade existente entre os fluxos para o canal TVI e RTP1 no final da noite.

□ **Período de Referência: a partir 22H/ 2Q**

Tendo como referência o período 22/2Q o decréscimo nos índices de duplicação apresenta uma evolução polinomial e não linear. Apenas num intervalo de meia hora existe um forte decréscimo nos seus valores (-30%). Considerando 23/1Q como referência as rectas *Tv_On* e *Tv_Off* cruzam-se às 24H. Apenas a partir da 1H existe uma maior proximidade entre as rectas *SIC* e *TV_On*. Os canais concorrentes pouco se diferenciam e os valores dos fluxos são inferiores aos 20% para a maior parte deste período. Por último considerando 0/1Q as rectas *Tv_On* e *Tv_Off* cruzam-se à 1H e perde-se o paralelismo existente entre as variáveis *SIC* e *Tv_On*.

O gráfico *F.4.5.III.7* representa os valores de fluxos entre os períodos adjacentes. O eixo temporal foi dividido em intervalos de 15 minutos, no entanto, apenas existem índices entre períodos associados ao início dos programas ainda que a nível parcial. Consequentemente, por vezes, os períodos adjacentes na realidade não o são no eixo temporal.

Verificamos que as maiores reduções em *Tv_On* estão associadas às 23H e simultaneamente verifica-se um maior distanciamento entre as rectas *SIC* e *TVI*. As maiores reduções em *Dup_SIC* correspondem ao final do espaço informativo e é acompanhada por um aumento em *Tv_Off* e em *TVI*. Nos períodos de maiores audiências para *Dup_SIC* existe uma maior proximidade de valores face a *TV_On* e uma redução muito significativa em todos os restantes valores dos fluxos de audiências. Segundo o gráfico que representa os valores dos fluxos distanciados de dois períodos verificamos que apenas a partir das 23H se verifica uma evolução mais desigual na variável *SIC* e *TV_On*. Os maiores valores destas duas rectas correspondem ao desenrolar do espaço informativo e à principal telenovela da noite.

IIIb. A evolução das decisões de ligar e desligar o televisor

O Gráfico *F.4.5.III.12* ilustra a evolução das variáveis *Tv_On* e *Tv_Off* para os períodos que decorrem entre as 19H e as 24H. A maior parte do painel opta por desligar o televisor entre as 23H e as 24H. Verificamos que o período de maiores audiências apresenta uma dimensão restrita para o canal *SIC*, podendo o mesmo ser prolongado através de uma correcta estruturação da grelha de programas, uma vez que até às 24H existe um acréscimo por volta dos 10% em *Tv_Off*. Obviamente que os factores orçamentais condicionam a estruturação da grelha de programas. Nos últimos anos este canal recorre à transmissão de telenovelas nacionais e às telenovelas da Rede Globo.

IIIc. Análises de variabilidade de índices de duplicação

Os gráficos F.4.5.III.8 a F.4.5.III.11 representam idênticas análises, mas considerando os intervalos de variabilidade.

O gráfico que ilustra a evolução dos fluxos a partir das 19H exemplifica a escassa variabilidade existente nos valores dos fluxos para o canal TVI e para a RTP1.

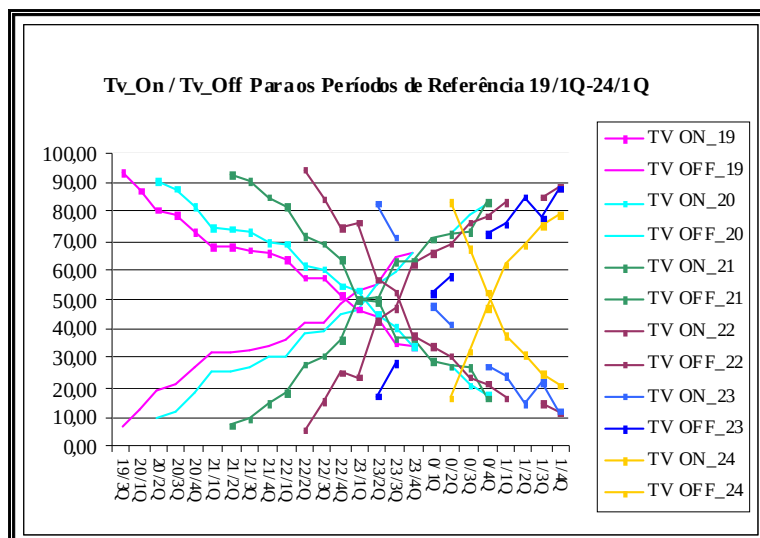


Gráfico F.4.5.III.12 – Evolução de Tv_On e Tv_Off

Para o período em análise verificam-se oscilações muito significativas nestes horizontes temporais, quer quanto aos horários de transmissão dos programas, quer nas tipologias e sequências em que são transmitidos. Corresponde essencialmente às faixas horárias que decorrem entre as 21/1Q e as 21/3Q. Portanto, a um período não muito extenso e que coincide, também, com as decisões de alguns telespectadores para desligar momentaneamente o televisor. As maiores amplitudes de valores para *Tv_On* e *Tv_Off*, também, correspondem a estes períodos. A partir das 22H existe uma nítida diferenciação entre os intervalos de variabilidade das variáveis *Tv_On* e *Tv_Off*.

Tomando como referência as 21/1Q as oscilações nos índices de duplicação são superiores. Os valores mais elevados correspondem a exposições ao canal RTP1 como consequência da transmissão dos eventos desportivos. Para 21/2Q a variabilidade também é significativa.

A partir das 22H os intervalos de valores reduzem-se consideravelmente. Os índices associados a *Dup_SIC* são superiores aos dos restantes canais até às 24H. A TVI neste contexto apresenta uma maior variabilidade de valores do que o canal estatal, o que é compreensível uma vez que ao nível de séries de ficção a RTP1 não tem uma grande repercussão nos índices dos canais concorrentes. As maiores exposições ao canal estatal, entre as 21H e as 22H, resultam apenas da transmissão de determinados eventos desportivos.

IIId. Índices de duplicação face ao espaço informativo da hora do almoço

As seguintes análises dizem respeito aos índices de duplicação face ao espaço informativo da hora do almoço, isto é, com 13/1Q como referência (Gráfico F.4.5.III.13). Com base nos valores médios observados existe paralelismo entre as rectas *Tv_On* e *Dup_SIC*.

Logo, existe uma certa independência entre as audiências deste canal face aos restantes canais. Os fluxos em geral são reduzidos. Apenas a RTP1 e a TVI apresentam valores ligeiramente superiores. A tendência ascendente dos valores de fluxos de audiências tem início por volta das 11H, seguindo uma evolução exponencial face a uma maior proximidade ao início do espaço informativo da tarde. A segunda parte do espaço informativo apresenta uma redução média no valor de 40% para SIC e de 20% para Tv_On.

O período 14/1Q apresenta uma redução de 20% face ao período anterior, 14/2Q uma redução de 10% para ambas as rectas e 15/1Q apresenta apenas um índice de duplicação médio face a 13/1Q no valor de 23% para SIC e de 32% para Tv_On. O início das 15H corresponde ao final do programa sucessor ao espaço informativo, como, por exemplo, desde 2007 às telenovelas da hora do almoço. O período 12/3Q apresenta uma percentagem de duplicação de 70% em Tv_On e de apenas 54% em SIC face ao início do espaço informativo da tarde.

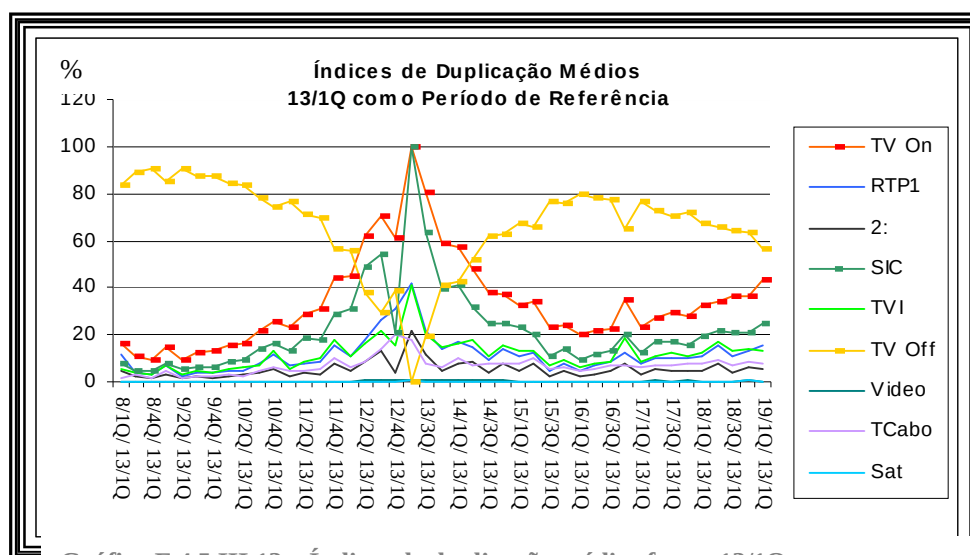


Gráfico F.4.5.III.13 – Índices de duplicação médios face a 13/1Q

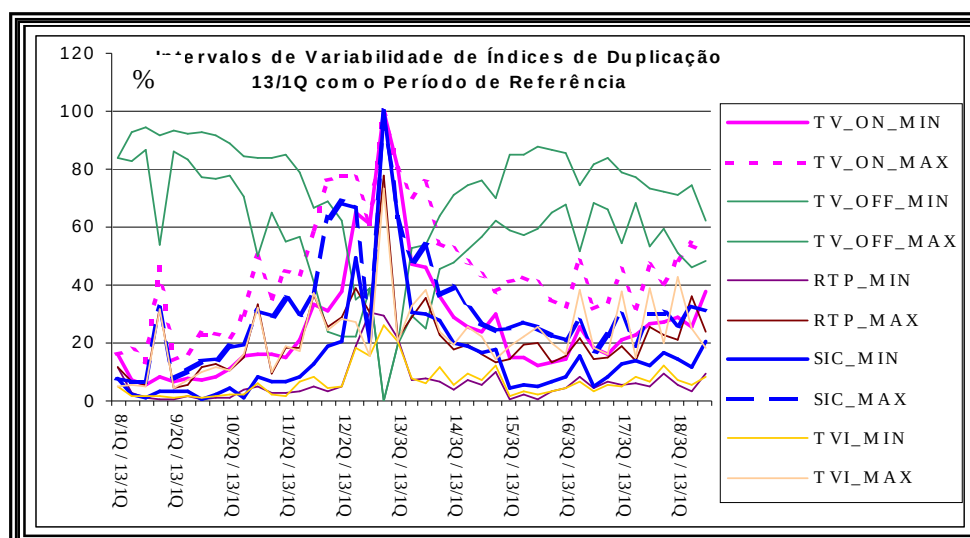


Gráfico F.4.5.III.14 – Variabilidade de índices de duplicação face a 13/1Q

IV. A influência dos factores estruturais nos índices de duplicação primários

Neste item pretendemos verificar quais são os factores estruturais que condicionam em maior grau os índices de duplicação primários entre programas ou faixas horárias directamente adjacentes, considerando a totalidade da emissão diária.

Na literatura da especialidade recorre-se em menor grau a esta tipologia de dados para representar a duplicação de audiências e consideram como variáveis explicativas apenas os valores de *ratings*, as tipologias e correspondentes sinergias e o número de programas na sua interface.

Nas análises a realizar consideram-se as seguintes variáveis explicativas:

(a) Índices de audiências do programa de referência e adjacentes.

Por exemplo, a saturação publicitaria, *Ratings*, *Shares*, *Taxas médias de audiências para o total da TV*, a fidelidade programática, durações, entre outros.

Os valores referem-se às partes programáticas envolvidas em cada par de observações e atendendo à globalidade da duração de ambos os programas.

(b) Variáveis de caracterização de cada um dos programas em cada par considerado.

Por exemplo, a duração dos programas global e parcial, tipologias programáticas de grau 1, grau 2 e grau 3; número de partes programáticas de cada um dos programas, número de intervalos no decorrer do programa global, períodos horários associados a ambos os programas, entre outras variáveis.

(c) Variáveis que inter-relacionam a programação:

Por exemplo, o quociente entre os valores de *ratings* parciais e globais de ambos os programas, o quociente entre os valores das taxas médias para o total da TV e dos *shares* de ambos os programas (globais e parciais), sinergias de tipologias de grau 1, 2, 3, ou seja, se coincidem ou não, a duração desde a conclusão do programa de referência e o início do programa sucessor, a duração desde o final do programa de referência e o início do programa sucessor, se ambas as partes programáticas pertencem aos mesmo programa, a tipologia da programação em contra-programação, se existem sinergias entre tipologias ou não, os valores de fluxos de audiências para com os canais concorrentes e as decisões de ligar ou desligar o televisor.

Recorre-se à Regressão Linear Múltipla e às Árvores de Regressão, por forma a comparar os resultados alcançados com as metodologias lineares e as abordagens não lineares. Na literatura da especialidade imperam as análises lineares e as análises incidem essencialmente nos períodos de maiores audiências.

Os seguintes resultados dizem respeito ao canal SIC.

➤ Do ponto de vista linear quais são as variáveis estruturais que se associam em maior grau com a duplicação primária das audiências? Em que medida as variáveis estruturais permitem prever este indicador do ponto de vista linear?

Atendendo à totalidade da emissão diária verificamos que os maiores valores de correlação face a *Dup_SIC* estão associados à variável *Tv_On* ($\beta=0,82$), *Tv_Cabo* ($\beta=-0,15$), *RTP2* (0,3), *Tv_Off* (-0,15) e *RTP1* (-0,09).

Um modelo de Regressão Linear incluindo as variáveis mencionadas apresenta um R^2 ajustado com o valor de 0,73 e apenas com *Tv_On* de 0,68 (calculado para uma amostra de teste com 30% das observações).

Webster (2006) considerando os *ratings*, as tipologias e número de opções disponíveis apenas consegue alcançar um coeficiente de determinação com o valor de 0,58 e em *Eastman, Newton, Riggs e Neal-Lundsford (1997)* este valor é de 0,56.

Logo, com os dados nacionais esta percentagem é superior e os fluxos de audiências para os canais concorrentes pouco acrescentam à variância explicada com base numa recta de Regressão que inclua apenas *Tv_On*, ou seja, a percentagem de telespectadores que permanece com o TV ligado entre os programas adjacentes.

Ainda assim, a existência de fluxos de audiência para os canais por TV Cabo penaliza os efeitos de herança observados no Canal SIC, assim como as decisões de desligar o televisor, ainda que a associação linear seja reduzida e muito inferior à da variável *Tv_On*. Os coeficientes associados aos canais estatais são muito reduzidos, com a RTP1 pontualmente a penalizar os fluxos entre os programas adjacentes e o Canal 2 a favorecer.

A variável *Tv_On* apresenta o maior valor do coeficiente beta, apresentando-se como a variável que do ponto de vista linear permite prever em maior grau a duplicação de audiências. Quanto maior for a percentagem da audiência que mantém o televisor ligado na sua interface maior será o valor da duplicação de audiências e existe alguma proporcionalidade na sua evolução. Com base em análises gráficas verificamos um certo paralelismo entre a evolução das duas rectas.

Existindo uma associação linear com *Tv_On* acaba por incutir alguma linearidade na relação entre os valores de *ratings* e as taxas médias de audiências para o total da TV. Logo, neste horizonte temporal existe para grande parte do painel lealdade para com o canal, e a duplicação de audiências resulta deste factor, encontrando-se fortemente dependente dos factores de disponibilidade e de exposição ao *media*.

Os programas que conseguem manter na sua transição um maior número de telespectadores acabam por apresentar maiores índices de duplicação.

Logo, existe alguma passividade nas selecções que, também, decorre da lealdade para com o canal e os programas. A lealdade ao canal acaba por ser construída com base no grau de envolvimento da população com a grelha de programação.

Verificamos que, por vezes, a principal telenovela da noite consegue valores de *ratings* que ultrapassam os valores do espaço informativo. Logo, também existe lealdade programática, mas que segundo as análises de lealdade programática não transcende os 50% do público que está presente em determinada transmissão de um programa.

Com excepção da variável *Tv_On*, a ausência da linearidade das relações existentes entre as variáveis é notória, sendo esta a metodologia mais utilizada na literatura da especialidade.

As restantes variáveis de planeamento e de caracterização da programação não apresentam coeficientes significativos, considerando a totalidade dos dados. No entanto, o planeamento e a estruturação da grelha de programas condiciona a evolução das variáveis *TV_on* e *TV_off*.

Os factores de disponibilidade e a exposição ao *media* são os principais factores que condicionam a duplicação das audiências a nível global, logo a previsão dos *ratings* encontra-se condicionada por parte da evolução das taxas médias de audiências para o total da TV que aparentemente, por vezes, apresenta alguma independência face à grelha de programação e aos conteúdos transmitidos.

Uma vez que os factores de disponibilidade condicionam fortemente os resultados, isto dificultará a previsão destes indicadores.

➤ Do ponto de vista linear quais são os valores de *ratings* que se associam em maior grau com a duplicação primária das audiências?

Na literatura da especialidade, frequentemente, utilizam-se os *ratings* para avaliar os valores de duplicação de audiências. Considerando os *ratings* programáticos globais e parciais, os valores de correlação linear entre a duplicação de audiências e os indicadores associados ao programa de referência e de *lead_in* são irrisórios para o horizonte temporal em análise.

Em relação aos valores de *ratings* do programa de *lead_out* existe uma maior associação do ponto de vista linear com o valor do *rating* parcial e com o valor de 0,21 e de 0,17 face ao valor do *rating* global sucessor.

Verificamos que existe uma maior associação com o valor parcial, uma vez que no decorrer de um programa poderá existir uma significativa diferenciação de valores. Apenas à hora do almoço e no horário nobre assistimos, por vezes, a uma maior estabilidade na evolução dos *ratings* no decorrer da duração global de um programa. Um modelo de RLM que incluía *Rat_ldin%* (*P*) ($\beta=1,22$), *Rat_Lout%* (*T*) ($\beta=-0,89$) e *Rat_ldin%* (*T*) ($\beta=-0,21$) apresenta um coeficiente de determinação de apenas 0,18.

Não existe associação linear entre as variáveis, mas os coeficientes de correlação linear de pearson são superiores para o programa de *lead_out*.

➤ Do ponto de vista não linear quais são as variáveis estruturais que se associam em maior grau com a duplicação das audiências?

Alternativamente, recorrendo a Árvores de Regressão e considerando a globalidade da emissão diária e das variáveis em análise, verificamos que a variável que condiciona os índices de duplicação observados também corresponde a *Tv_On*. Os padrões de disponibilidade da audiência condicionam a duplicação de audiências e, conseqüentemente os *ratings*.

Se $Tv\_On < 77,35\%$ existe um valor de duplicação médio de 45% (diferencial superior a 30% face a *Tv_On*); para $Tv\_On \in [77,3\% - 92,2\%]$ o valor de duplicação médio é de 65% e para $Tv\_On > 92\%$ o valor é de 84%. Portanto, grande parte dos telespectadores, no geral, mantêm-se no canal SIC de programa para programa, sendo a nível relativo superior para os maiores níveis de *Tv_On*. Os programas que mantêm um menor número de telespectadores com o televisor ligado na sua interface são também os que perdem um maior número de pessoas para os canais concorrentes ($Tv\_On < 77,35\%$), o que para o período em análise, atendendo à evolução dos *ratings*, corresponde em maior grau a algumas telenovelas de ficção nacional do canal TVI.

Logo, existe alguma fidelização por parte dos telespectadores a este canal, optando a partir de determinadas faixas horárias por desligar o televisor em detrimento de mudanças de canal. Predomina a lealdade ao canal nos períodos de maiores audiências, existindo em faixas horárias mais tardias perante quem apresenta uma maior disponibilidade e envolvimento com o meio televisivo uma transição de canal, imperando a lealdade programática, uma vez que nesses períodos o canal SIC, também, transmite telenovelas, mas essencialmente brasileiras da Rede Globo.

Se existisse variação ao nível das tipologias programáticas entre o canal SIC e a TVI seria consequência da lealdade tipológica, mas uma vez que estas coincidem, os fluxos de audiências devem estar associados à lealdade programática.

No horizonte temporal em análise existe um crescente interesse pela ficção nacional e, por vezes, atendendo ao sucesso da principal telenovela da SIC, a TVI reserva os seus melhores programas para as faixas horárias mais tardias onde mais facilmente consegue retirar audiência ao seu canal concorrente. No decorrer dos anos, por vezes, invertem-se estes papéis e a SIC acaba por ser penalizada por parte das séries de ficção da TVI transmitidas no início da noite e beneficiando da popularidade que as telenovelas da Globo no final da noite apresentam para determinados segmentos de telespectadores. Os conteúdos certamente que influenciam os fluxos de audiências, assim como as variáveis estruturais, nomeadamente a lealdade.

Para os valores mais elevados de *Tv_On* (*Tv_On*>92%), frequentemente, apenas a exposição aos canais por TV Cabo condiciona negativamente a duplicação das audiências.

Logo, quem selecciona este canal em alguns dos anos considerados adere em maior grau aos conteúdos mais específicos dos canais por TV cabo e, não, aos canais generalistas concorrentes. Esta situação posteriormente alterou-se com a crescente popularidade dos programas da TVI.

Para valores de *Tv_On* $\in$ [77,3%-92,2%] a exposição ao Canal 2 favorece este indicador.

Em faixas horárias com uma exposição televisiva ainda muito significativa, os telespectadores, por vezes, sintonizam o Canal 2, mas com uma duração que em vez de penalizar o canal SIC inclusive se associa positivamente com a duplicação de audiências observada neste canal. Uma vez que parte do painel não apresenta TV Cabo e face à procura dos espaços informativos, nestas faixas horárias existe alguma exposição ao Canal 2, mas insuficiente para penalizar a duplicação de audiências do canal SIC.

Para os menores valores de *Tv_On* (<77,35%) a variável *Tv_Off* penaliza este indicador: (se *Tv_Off*<31% a *Dup_SIC_Medio*=0,51; se *Tv_Off* >31% e <37% a *Dup_SIC_Medio*=0,43 e 0,36 caso contrário). Em função da ordem de grandeza da variável *Tv_Off* o grau de duplicação de audiências vai variando.

Os canais generalistas parecem não influenciar o público deste canal, no período que apresenta uma crescente consolidação da posição do canal TVI. Neste período o canal TVI que pode ser visto como um concorrente da SIC, não apresenta frequentemente coeficientes estatisticamente significativos, quer do ponto de vista linear, quer do ponto de vista não linear para a maior parte da emissão diária. Apenas em horários mais tardios assistimos a um fluxo de audiências entre estes dois canais. Para os restantes anos, os resultados alteram-se existe uma maior concorrência nos períodos de maiores audiências.

Para os menores valores de *TV_On*, os valores de duplicação acabam por ser condicionados em maior grau pelo número de telespectadores que decide desligar o televisor do que por parte dos fluxos de audiências para os canais concorrentes. Em períodos com um maior número de telespectadores a programação por TV Cabo poderá retirar níveis de audiências a este canal; no seguinte patamar de exposição o Canal 2 perde audiência relativamente a este canal e em períodos de menor exposição os padrões de disponibilidade condicionam a duplicação de audiências em detrimento dos fluxos para os canais concorrentes.

Considerando a globalidade dos dados, as decisões de ligar ou desligar o televisor e, consequentemente os valores das taxas médias de audiências para o total da TV na interface de dois programas são as variáveis que condicionam os índices de duplicação e por conseguinte os valores dos *ratings* programáticos.

Alterações ao nível do desempenho e imagem do canal podem influenciar estes resultados, mas aparentemente existe uma identificação do público com determinados canais não sendo fácil retirar público dos canais concorrentes, pelo menos a curto prazo.

Nas análises realizadas aos valores de *ratings* verificamos que apenas nos contextos em que são transmitidos programas que geram observações atípicas poderá existir uma maior repercussão nos *ratings* dos canais concorrentes. Inclusive inúmeros jogos de futebol transmitidos nos canais generalistas estatais não conseguem alterar a evolução dos *ratings* ao nível dos canais privados.

➤ Qual a influência das variáveis temporais na duplicação das audiências?

Atendendo às Árvores de Regressão e considerando as variáveis temporais e as taxas médias de audiências para o total da TV, a variável que se associa em maior grau com os índices de duplicação primários programáticos são a Hora_Prog (J), ou seja, a hora do programa de referência face ao qual se calcularam as taxas de duplicação condicionadas e primárias. Os mais valores de duplicação estão associados a algumas horas que se repartem no decorrer de todo o dia com uma Dup_Média=80%: 9/4Q (representa o período que decorrer entre as 9.45 e as 10H), 11/1Q, 12/1Q a 12/3Q, 13/4Q, 15/3Q, 19/1Q-19/4Q, 21/1Q, 21/2Q, 2/1Q e 2/2Q.

Não engloba a maior parte dos períodos aos quais podemos associar maiores valores de *ratings*. Os valores dos *ratings* podem resultar de um menor número de telespectadores, mas que assistem aos programas durante um maior intervalo de tempo ou de um maior número de indivíduos, mas com uma exposição mais breve, uma vez que reflecte a taxa média de audiência por minuto.

A duplicação de audiências deverá ser, obviamente, mais elevada na primeira situação, quando existe um maior envolvimento com a programação. Logo, se nos períodos de maiores audiências os *ratings* resultam de um maior número de exposições, mas mais breves a duplicação de audiências não terá que ser obrigatoriamente superior.

Na literatura da especialidade, frequentemente, associam-se maiores valores de efeitos de herança a maiores valores de *ratings*, mas em geral as análises, também, incidem somente nos períodos de maiores audiências.

Os programas sucessores ao espaço informativo da noite, e que em geral são breves, apresentam maiores valores, uma vez que a audiência poderá esperar pelas telenovelas do início da noite, assim como, no pré-prime-time, uma vez que a audiência espera pelo início do espaço informativo da noite. Nestes períodos horários os valores observados são função dos valores de *Tvr_Prog%* totais e parciais.

Os valores mais reduzidos das taxas médias de audiências para o total da TV beneficiam os efeitos de herança entre os programas (com $Tvr\_Prog\%\_J < 4,45\% \rightarrow Dup\_Média = 95\%$ e no valor de 77% caso contrário).

Os programas no decorrer dos quais a audiência média para o total da TV é menor associam-se a uma maior permanência do público no seu decorrer (uma audiência de menor dimensão, mas mais leal à grelha de programas). Para os maiores valores de $Tvr\_Prog\%(T)_J$ a duplicação de audiências será beneficiada por maiores valores parciais das taxas médias de audiências para o total da TV, (se $Tvr\_Prog\%(P)_J < 5\%$, $Dup\_Média = 55\%$ e no valor de 78% caso contrário).

A duplicação de audiências encontra-se condicionada pela inter-relação existente entre as taxas médias de audiências para o total da TV associadas às partes programáticas envolvidas.

Em períodos de maiores níveis de audiências quanto maior for a audiência total disponível no programa de referência mais elevados, também, serão os índices de duplicação. Os menores valores de índices de duplicação correspondem essencialmente aos suportes transmitidos entre as 8H e as 10H da manhã, ao início da tarde (13/4Q-15/1Q) ou entre as 23H e as 2H da madrugada (com uma $Dup\_Média = 47\%$).

No início da manhã durante o pequeno almoço os telespectadores podem assistir pontualmente à TV, procurando os conteúdos informativos. Logicamente a exposição será breve. O mesmo sucede após o almoço, tentando a SIC cativar as audiências, dia após dia, transmitindo cada vez mais telenovelas da TV Globo. Por volta das 23H verifica-se um significativo acréscimo nos valores de Tv_Off e, logo a duplicação de audiências, também, apresenta valores que são reduzidos.

O período da manhã apresenta os maiores ($Dup\_Média = 58\%$) valores seguindo-se a noite ($Dup\_Média = 49\%$) e os restantes períodos horários encontram-se interrelacionados com um valor médio de duplicação das audiências equivalente a 40%. A um nível intermédio ($Dpr\_Médio = 63\%$) os períodos horários, também, condicionam os valores da duplicação das audiências. Segundo as Árvores de Regressão, os maiores valores dos índices de duplicação entre programas adjacentes associam-se à manhã, ao almoço e à tarde; segue-se o horário nobre, o princípio tarde e a noite e com menores valores apenas o final da noite.

Realizando uma análise ANOVA e considerando como factor apenas o período horário verificamos que os índices de duplicação a partir do pré-prime-time vão sucessivamente decrescendo em valor (o período da noite com $Dup\_Médio = 55\%$ e o final da noite com $Dup\_Médio = 45\%$). Desde o almoço até ao horário nobre os valores não se diferenciam significativamente.

A manhã apresenta valores da mesma ordem de grandeza aos observados no pré-prime-time (*Dup_Médio*=70%). As faixas horárias do almoço e da tarde equiparam-se ao horário nobre com uma *Dup_Média*=65%.

Realizando, uma análise ANOVA a II factores que permita a avaliação da influência do período semanal ou de fim-de-semana nos valores de duplicação associados aos períodos horários verificamos que estes não se diferenciam. O efeito principal de considerarmos dias laborais ou de fim-de-semana não influencia a dimensão destes índices.

Considerando a média de valores apenas no decorrer da tarde e ao fim-de-semana a duplicação de audiências apresenta um diferencial negativo no valor de 3-4%. Logo, os períodos de lazer não levam a uma maior duplicação de audiências, nem mesmo nos períodos da noite.

➤ Qual a influência das tipologias programáticas de grau 1, 2 e 3 na duplicação das audiências?

Atendendo às tipologias programáticas de grau I para o período em análise algumas transmissões de desporto e programas de cultural geral ou conhecimento minimizam estes valores (*Dup_Média*=40%).

Os programas destinados aos mais jovens apresentam os maiores valores de duplicação com uma *Dup_Média*=73%, assim como os programas que aparecem com a designação

tipológica de Diversos e que correspondem a programas de muito curta duração intercalados entre os espaços informativos e as principais telenovelas da noite, como, por exemplo, os Boletins meteorológicos (*Dup_Média*=85%). As restantes tipologias não se diferenciam entre si.

Considerando as tipologias de grau II, os programas de apoio à comunidade, os musicais e os passatempos maximizam este índices enquanto que os magazines, as mini-séries, os telefilmes e as cerimónias de desporto minimizam-nos (*Dup_Média*=40%).

Em relação às tipologias de grau III verificamos que as reportagens e os sketches humorísticos beneficiam estes índices, mas a informação cultural penaliza-os. No que diz respeito aos filmes apenas a tipologia de romance apresenta maiores índices de duplicação, enquanto que a ficção científica, a acção e o suspense minimiza-os.

Regra geral, ao sexo feminino associamos uma maior tendência para a duplicação de audiências, daí a associação que se verifica entre as tipologias programáticas e a duplicação das audiências.

Os espaços informativos apresentam uma duplicação de audiências que se insere no nível médio observado. Através das Árvores de Regressão verificamos que a tipologia de grau I condiciona os valores dos índices de duplicação: os valores mais reduzidos associam-se à cultural geral e às cerimónias de desporto (*Dup_Média*=43%), segue-se o

divertimento, a ficção e a informação (*Dup_Média*=64%) e com maiores valores os programas de Juventude (*Dup_Média*=83%).

Para as tipologias de Divertimento, Ficção e Informação, a variável que as diferencia são as tipologias de grau II: musicais (*Dup_Média*=75%); talk-shows, noticiários, séries, telenovelas e filmes (*Dup_Média*=63%); humor (*Dup_Média*=65%); drama e ficção científica (*Dup_Média*=36%); suspense, stand up comedy e beleza (*Dup_Média*=53%).

➤ Quais as variáveis de planeamento que condicionam a duplicação de audiências?

Uma vez que considerando as faixas horárias e os fluxos de audiências entre os canais estas são as variáveis que condicionam em maior grau a duplicação de audiências, nas seguintes análises estas foram eliminadas com excepção de *Tv_On* e *Tv_Off*.

Considerando a globalidade dos restantes indicadores de audimetria, os factores de planeamento, as variáveis de caracterização da programação e do seu desempenho verificamos que a duplicação de audiências será em primeiro lugar função da evolução de *Tv_On*.

(a) Para os valores de *Tv_On* superiores aos 92% a fidelidade programática do programa de referência condiciona estes valores (*Dup_Média*=85%).

Maiores valores de fidelidade programática levam a uma maior duplicação de audiências. Quanto maior for o tempo de exposição no decorrer de um programa o que denota um maior envolvimento com a programação, maior será, também, a duplicação face ao programa sucessor.

(a1) Para os valores de fidelidade programática superiores aos 95% a duplicação apresenta-se como uma função dos níveis de *Tv_On*;

(a2) Para os valores de fidelidade no seguinte intervalo (68%-95%) a duplicação das audiências também será uma função de *Tv_On*;

(a3) Para os valores de fidelidade inferiores aos 68% a duplicação das audiências será função de *Rat_Prog%(T)_I* ou *Tvr_Prog%(T)_I* que representa o *rating* e a taxa média de audiência para o total da TV (global) do programa sucessor.

Logo, para os maiores valores de fidelidade as decisões de manter o televisor ligado na interface de dois programas condicionam a ordem de grandeza da duplicação de audiências. Para os menores valores de fidelidade face a um programa a duplicação de audiências depende dos indicadores associados ao programa subsequente, nomeadamente dos seus *ratings* e das taxas médias de audiências para o total da TV. Face a um programa menos apelativo será o interesse, a disponibilidade e a faixa horária do programa subsequente que condiciona a duplicação de audiências.

(b) Para valores de *Tv_On* no intervalo 77% a 92% a fidelidade programática do programa de referência condiciona estes valores e, posteriormente a duração do programa sucessor com uma influência positiva.

(c) Para os menores valores de *Tv_On* será a ordem de grandeza dos valores de *Tv_On* ou de *Tv_Off* que condiciona a evolução da duplicação das audiências, assim como os *shares* dos programas sucessores com uma influência positiva.

Apenas a fidelidade programática condiciona em maior grau a duplicação das audiências. Todos os restantes indicadores de audimetria totais ou parciais, assim como as durações programáticas e o número de partes que integram os programas, entre outras variáveis, não se apresentam com coeficientes estatisticamente significativos. Será necessário seleccionar variável a variável para avaliar a sua repercussão na duplicação das audiências.

Por exemplo, a cada par de partes programáticas adjacentes consideradas associa-se uma variável nominal que traduz se ambas pertencem ao mesmo programa. Verificamos que o diferencial médio de valores é pouco significativo, com o valor falso a maximizar ligeiramente a duplicação de audiências. A duplicação de audiências não é superior se consideramos partes programáticas associadas ao mesmo programa.

Considerando o número de partes programáticas que integram o programa de referência o coeficiente de correlação linear não é significativo. O coeficiente de correlação linear entre a duplicação de audiências e a variável que representa o quociente entre as durações programáticas de cada um das partes programáticas é de 0,34 e será de 0,1 entre as durações globais.

Considerando as durações dos programas envolvidos, a duração do programa de referência apresenta um coeficiente de correlação linear de pearson com o valor de 0,67 e a duração do programa sucessor um valor de 0,45. Uma maior duração programática relativa de *lead_out* está associada a maiores valores de duplicação de audiências. Uma maior duração programática do programa de *lead_in* penaliza a duplicação de audiências face ao programa sucessor, enquanto que a duração de *lead_out* está associada a maiores valores.

Considerando o dia da semana, a Sexta, o Domingo e a Segunda maximizam os valores de duplicação face aos restantes dias da semana, sendo estes os dias que apresentam maiores valores no decorrer da totalidade da semana. O Sábado e a Quinta por sua vez correspondem aos dias que minimizam estes indicadores para o período em análise.

□ **A Influência dos factores estruturais nos índices de duplicação primários adjacentes (Base de dados que inclui os suportes entre as 19H-24H)**

➤ Do ponto de vista linear quais são as variáveis que se associam em maior grau com a duplicação das audiências nos períodos de maiores audiências?

Considerando como variável dependente a duplicação de audiências primária e como variáveis explicativas os fluxos de audiências entre os canais e as decisões de ligar e desligar o TV verificamos que um modelo de Regressão Linear Múltipla com apenas *Tv_On* apresenta um coeficiente de determinação no valor de 0,60 ($\beta=0,76$).

Adicionando o Canal 2 e a RTP1, apresentando este último um coeficiente negativo uma vez que incluímos a totalidade da programação (inclui os eventos desportivos), o coeficiente de determinação é de 0,69. Incluindo a variável TVI ($\beta=-0,27$) o coeficiente é de 0,72. O vídeo, a TV Cabo e os canais por satélite não apresentam coeficientes significativos. Atendendo a estes resultados assim como, se verificou para as análises considerando a totalidade da emissão diária, os efeitos de herança estão essencialmente condicionados pelas decisões de ligar o televisor, ou seja, pelas taxas médias de audiências para o total da TV.

A capacidade explicativa da exposição aos canais concorrentes é muito menor. Nas análises exploratórias constatou-se que apenas programas com um maior desempenho dos canais concorrentes conseguem influenciar a evolução dos *ratings* do canal SIC. A TVI e a RTP1 apresentam uma influência negativa e o Canal_2 positiva, ou seja, maiores exposições a este canal estão também associadas a maiores exposições ao canal SIC.

Considerando os valores de fluxos de audiências do ponto de vista linear conseguimos para uma amostra de teste com 30% das observações um $R^2=0,67$.

	Coeficientes de Correlação Parciais	Coeficientes de Correlação Totais
Dup. Primária Baixa (<60%)	0,87	0,6
Dup. Primária Alta (>60%)	-0,18	0,87
	Rat_P_I	Rat_T_I
Dup. Baixa	0,3	0,1
	Rat_P_J	Rat_T_J
	0,18	0,18
Dup. Alta	Rat_P_I	Rat_T_I
	0,43	0,09
	Rat_P_J	Rat_T_J
	0,15	0,04

Tabela F.4.5.IV.1 – Correlação entre *ratings* de pares de programas adjacentes sob elevada e reduzida duplicação primária de audiências

Considerando apenas a associação com os valores dos *ratings* o maior valor corresponde ao *rating* parcial do programa de *lead_out* ($\beta=0,55$ e $R^2=0,30$). Adicionando o *rating* parcial do programa de referência ou de *lead in* o $\beta=-0,52$ e o $R^2=0,43$. Não existe associação com os valores globais de *ratings*. Logo, a associação com o valor sucessor é positiva e com o antecessor negativa.

Considerando os valores de *ratings* e de taxas médias de audiências para o total da TV, assim como, os fluxos de audiências entre os canais com a variável *TV_On* o coeficiente de determinação é de 0,58. Incluindo também a variável Canal 2 com um coeficiente positivo e *Tv_Off* com um coeficiente negativo o coeficiente de determinação ascende a 70%. Incluindo também as variáveis *TVI*, *Rat_Prog%(P)_J*, *TV_Cabo*, *Tvr_Prog%(P)_I* com coeficientes negativos e *TVR_Prog%(P)_J*, *Rat_Prog%(T)_J* e *Rat_Prog%(P)_I* com coeficientes positivos o coeficiente de determinação é de 0,80 o que representa um acréscimo de 12% face a considerar apenas *Tv_On*.

O sinal associado às taxas médias de audiências para o total da TV não coincide com o dos valores de *ratings*. Enquanto que os valores de *ratings* parciais de *Lead_in* se associam negativamente com as duplicações primárias os valores das taxas médias de audiências para o total da TV de *lead_in* são positivos. Incluindo a totalidade das variáveis, as restantes variáveis e indicadores de audimetria não apresentam coeficientes estatisticamente significativos.

➤ Em que medida a correlação entre pares de *Ratings* traduz a existência de uma duplicação primária alta ou baixa e considerando apenas os dados a partir das 19H?

Como consta da *Tabela F.4.5.IV.1* para pares de programas adjacentes que apresentam índices de duplicação primária superiores aos 60%, os coeficientes de correlação considerando os valores de *ratings* parciais é de -0,18 e no valor de 0,87 considerando os *ratings* globais.

Para os valores de duplicação primária inferiores aos 60%, o grau de correlação entre os *ratings* parciais é de 0,87 e de 0,6 para os valores de *ratings* totais.

Logo, as situações com uma menor duplicação de audiências são na realidade as que apresentam maiores associações lineares entre os *ratings* de observações adjacentes. Uma maior correlação entre os *ratings* de observações adjacentes não nos permite concluir que a duplicação de audiências poderá ser considerável.

Para os maiores valores de duplicação, as variáveis inclusive relacionam-se de forma inversa, ainda que a associação linear seja pouco significativa. Por exemplo, em períodos com uma tendência descendente ou ascendente de valores *ratings* a duplicação das audiências ainda poderá ser razoável.

Não se confirmam os resultados de Webster (2006), segundo o qual para valores de duplicação superiores aos 60% a correlação é de 0,84 e para o conjunto complementar de 0,82.

Recorrendo aos dados de duplicação de audiências em vez de utilizar os dados de duplicação primária, também verificou que os contextos de menor duplicação apresentam uma maior correlação (0,82 e de 0,56 quando a duplicação é superior).

➤ Qual a repercussão de factores qualitativos, no grau de duplicação de audiências?

Considerando a influência da variável nominal que representa se cada parte corresponde ao mesmo programa, o falso apresenta um maior valor médio de duplicação primária.

A duplicação primária não é superior quando cada par de partes programáticas considerada corresponde ao mesmo programa.

Considerando a tipologia programática do programa de referência de grau 1 verificamos que as tipologias de “Diversos” e de “Outros” correspondem às que maximizam este indicador e incluem programas sucessores dos espaços informativos e em geral com uma duração reduzida.

Entre as tipologias que se associam em maior grau aos programas, como, por exemplo, a Informação, o Divertimento e a Ficção não se observam diferenças nas médias associadas à duplicação de audiências (valor médio de 60%

para a duplicação primária). O Desporto minimiza estes valores e apresenta uma grande amplitude de valores de duplicação.

A duplicação de audiências entre suportes transmitidos durante a semana é superior face ao suportes transmitidos ao fim-de-semana (*Dup_Média*=66% para a semana vs. *Dup_Média*=59% ao fim-de-semana).

Também verificamos que ao Domingo os valores são inferiores aos valores do Sábado (*Dup_Média*=55% ao Domingo e *Dup_Média*=63% ao Sábado) e no decorrer da semana a Sexta e a Segunda diferenciam-se ligeiramente com maiores valores de duplicação das audiências.

Considerando a variável que representa o período horário apenas a noite se diferencia com menores valores (*Dup_Média* =52%).

Entre o pré-prime-time e o prime-time as médias observadas coincidem nos 70%.

➤ Do ponto de vista não linear quais são as variáveis que se associam em maior grau com a duplicação das audiências nos períodos de maiores audiências?

Considerando a globalidade dos dados a identificação dos programas corresponde à variável que se associa em maior grau com a variável dependente. Os programas do início da noite apresentam maiores valores de duplicação (*Dup_Média*=70% vs. *Dup_Média*=49%, caso contrário).

O espaço informativo da noite no decorrer de toda a semana e os seus programas sucessores, incluindo a principal telenovela da noite maximizam a duplicação de audiências.

Os programas que minimizam a duplicação de audiências apresentam valores superiores quando a fidelidade programática supera os 40% (duplicação primária média entre os 37% e os 54%). Os programas que maximizam a duplicação de audiências apresentam valores superiores para as menores durações programáticas do programa de referência.

Excluindo, a identificação do programa, a variável *TV_On* surge como a variável que condiciona a duplicação de audiências. Esta Árvore de Regressão apresenta uma percentagem de variância explicada no valor de 86%.

Se *TV_On* for superior aos 83% a duplicação de audiências média corresponde a 73% e ao valor de 52% caso contrário.

Os maiores valores de duplicação encontram-se condicionados e favorecidos pelo quociente entre os *ratings* totais de ambos os programas. Quanto maior for a proximidade de valores de *ratings* entre os programas adjacentes maiores será também a duplicação de audiências (*Dup_Média*=74% vs. *Dup_Média*=68%). Para os menores valores de duplicação de audiências, a fidelidade programática favorece este indicador. Para os maiores valores de fidelidade a exposição ao canal RTP1 favorece este indicador.

Considerando *Tv_On* e *Tv_Off*, assim como os fluxos para canais concorrentes apenas as decisões de ligar e desligar o televisor condicionam a ordem de grandeza da duplicação de audiências (83% de variância explicada). Atendendo aos resultados para avaliar a influência dos factores de planeamento será necessário considerar metodologias de *Temporal Reasoning*.

V. As necessidades de análise para realizar a previsão de *Ratings* em períodos que apresentem uma maior variabilidade

Em função das faixas horárias e dos contextos em análise, diferentes abordagens e metodologias são necessárias para conseguir alcançar a futura previsão de valores de *ratings*.

As taxas médias de audiências ponderam quem selecciona um programa quanto à duração do seu contacto. Atendendo às correlações existentes entre os valores dos *ratings* de faixas horárias adjacentes, as regressões locais e as regressões polinomiais parecem ser uma boa opção para tentar prever os valores de *ratings* nos períodos de maiores audiências.

Alternativamente, e como foi referido no capítulo VI, no limite seria necessário quantificar diversas situações e que se resumem em quantificar quem poderá seleccionar determinado programa em determinado dia, semana e hora, o qual poderá variar nas faixas horárias em que é transmitido e posteriormente, quantificar a duração dessa mesma exposição. Devemos considerar os diversos contextos de duplicação das audiências.

Para cada suporte uma percentagem da audiência será herdada de episódios anteriores, ou seja, dos comportamentos de “repeat viewing” que também podem apresentar alguma ciclicidade, mas considerando suportes com algum afastamento temporal.

Atendendo às análises realizadas aos índices de lealdade programática estes valores são no máximo de 50% e não se observa uma forte repercussão de um maior afastamento temporal entre os suportes publicitários.

Para esta fracção do painel, em função das selecções passadas podemos calcular uma probabilidade de selecção e estimar a duração da exposição média com base no seu histórico de valores, considerando um longo horizonte temporal, uma vez que os comportamentos não apresentam ciclos de curta duração. Nestes contextos podemos recorrer a modelos binomias logísticos e à classificação de Bayes.

Contudo, para uma elevada percentagem do painel que não apresenta uma exposição mais regular ao meio televisivo será muito difícil prever quando efectivamente irá ligar o televisor e as respectivas selecções. Consequentemente, devem ser adoptadas perspectivas de análise alternativas e que se descrevem no Capítulo VI.4.5.

Se nos restringirmos à fracção do painel que apresenta uma exposição mais regular, será necessário quantificar a duplicação de audiências entre os episódios e uma outra componente relativa aos efeitos de herança que, por sua vez se podem interseccionar ou não ou resultarem dos fluxos de audiências por parte de canais concorrentes.

Face a quem poderá seleccionar determinado programa ou faixas horárias será necessário estimar durante quanto tempo para proceder ao cálculo das taxas médias de audiências que ponderam os telespectadores em função da duração da exposição.

Nestes contextos as perspectivas mais adequadas de construção de planos de *media*, será através da construção de Regras de associação e recorrendo ao *Data Mining* Sequencial.

Também é necessário verificar quais são as alterações que se verificam quando são transmitidos eventos desportivos, os quais apresentam repercussões muito diferenciadas na evolução dos *ratings*.

Em síntese, apresentam-se as seguintes propostas de análise:

□ Regras de Associação e com dados dos ficheiros de base de audimetria

Para cada Dia *i* e para cada telespectador *i* no decorrer da totalidade da emissão identificar as horas de ligar, desligar e mudar de canal.

Por Exemplo: Ind. *I* <Dia *i*, 12H Tv_On e SIC, 14H TVI, 16H Tv_Off> e Ind.1 < Dia *i*, 19H, Tv_On e SIC, 21.30H TVI, 22H RTP, 23.30H Tv_Off>

A partir dos dados dos ficheiros de base de audimetria podemos construir uma base de dados onde a cada Telespectador *I* se associa <Dia *i*, Prog *X*, Prog *Y*, Prog *Z*> considerando os programas em cada sessão para os quais assiste a pelos menos *X*% da sua duração, ou, independentemente da duração da exposição.

A partir destes dados podemos verificar quais são os programas que apresentam uma maior probabilidade de serem seleccionados em conjunto em cada sessão televisiva para cada telespectador e no decorrer da semana.

Posteriormente, verificamos quais são os canais a que pertencem e tipologias, avaliando a existência de comportamentos de lealdade programática, tipológica e ao canal, assim como a propensão para os efeitos de herança, seleccionando ou não programas frequentemente adjacentes.

Apenas verificamos quais são os programas seleccionados em conjunto com uma maior frequência, atendendo à duração da exposição seleccionada. Podemos restringir as análises aos programas que são vistos na sua totalidade, contribuindo a sua duração global para o cálculo da taxa média de audiência.

Em relação à duração da exposição também podemos considerar 50% da duração do programa, sendo esse o contributo para as taxas médias de audiências. Se seleccionarmos apenas 1 segundo seria necessário para cada programa analisar o seu historial de exposições considerando um longo horizonte temporal e calcular um valor médio ou intervalo de variabilidade.

Se seleccionarmos $x\%$ da sua duração podemos assumir essa duração como um valor médio de contacto com o programa. As análises podem ser realizadas em função do dia de semana.

Esta abordagem apresenta a desvantagem de ser difícil identificar sequências programáticas mais habituais em virtude das alterações na programação e das frequentes oscilações horárias.

Por exemplo, face à grelha de programação em 2010 onde a TVI e a SIC mantêm alguma regularidade horária e programática no decorrer da semana, esta opção seria viável. Com os dados de alguns dos anos considerados, como por exemplo, em 2004 será mais questionável a sua validade.

Alternativamente, realizar as análises em função das faixas horárias, por exemplo, para períodos de 30 minutos ou menores: <Ind: I, 19-19.30 SIC, 20-20.30H, TVI, 20.30-21H Tv_Off>. Verificar quais as sequências de canais seleccionados para as faixas horárias pré-definidas.

Verificamos qual a probabilidade de ter o televisor ligado e sintonizado em determinado canal. Em relação ao tempo de contacto, uma vez mais seleccionar $x\%$ e considerar esse valor como médio ou calcular um intervalo de variabilidade e aplicar metodologias simbólicas.

Em qualquer uma destas opções poderá facilitar a realização das análises proceder ao clustering dos telespectadores e calcular indicadores agregados.

□ **Regras Sequenciais**

As Regras de Associação não têm em atenção a dimensão temporal, apenas associações entre programas, logo em alternativa podemos recorrer a Regras Sequenciais. Contudo, segundo as análises de duplicação realizadas, a duplicação entre segmentos horários que não sejam próximos apresenta reduzidos valores.

□ **Análises Logit**

Para o cálculo da probabilidade de um telespectador seleccionar um programa, atendendo ao historial das suas selecções ou canal em função das diversas faixas horárias podemos recorrer aos métodos bayesianos para avaliar a repercussão de determinados factores estruturais e às análises Binomiais Logit para caracterizar o perfil sociodemográfico e, não só de quem apresenta maior probabilidade de selecção.

□ **Avaliação de comportamentos de lealdade**

Podem ser avaliados recorrendo às Regras de Associação e Sequenciais em função de cada telespectador ou por segmentos. Alternativamente recorrer a métodos factoriais e de MDS complementadas com análises de regressão, clustering e de classificação em contexto simbólico e clássico.

No seguinte apêndice constam algumas das análises realizadas neste contexto. O *Data Mining* não proposional será necessário para a construção dos planos de *media* quando se consideram os segmentos com uma exposição mais variável e neste contexto será especialmente relevante as Regras de Associação, os Métodos Bayesianos e o *Data Mining* Multirrelacional.

F.5. A SELECÇÃO DOS SUPORTES A INCLUIR EM PLANOS DE *MEDIA*

F.5.1. Introdução

A exposição televisiva tem sido analisada segundo diversas perspectivas, frequentemente, contraditórias (*Clancy, 1994*):

Four decades of research have demonstrated that people watch TV in different ways. Television – viewing behaviour has been explained in terms of user’s gratifications, instrumental and ritualized used, audience availability and audience duplication. The variable use of TV makes audience measurement an especially difficult task. (*p.1*)

As múltiplas dimensões que caracterizam o consumo televisivo contribuem para uma maior dificuldade quando se pretende a sua quantificação e previsão. Quando se constroem os planos de *media* será necessário seleccionar os suportes mais adequados ao público-alvo pretendido, independentemente do processo de decisão que lhe é na realidade subjacente. Deverá existir alguma duplicação de audiências entre os suportes seleccionados, os quais devem corresponder aos que apresentam uma maior “lealdade” ou adesão por parte do seu público. As análises de duplicação de audiências que constam dos dois apêndices anteriores contribuem, conseqüentemente, para a selecção dos suportes e relacionam-se com a previsão das audiências (*Fernandes, 2001*):

No âmbito da duplicação das audiências, pode-se, ainda referir duas outras características da duplicação da audiência: a lealdade ao canal e a repetição. Tal como o efeito de herança é uma previsão, a lealdade ao canal assume a mesma forma, ou seja, tem que ver com o facto das pessoas continuarem a assistir a um determinado programa mesmo que emitido em dias diferentes.

A análise da duplicação das audiências insere-se essencialmente na perspectiva estrutural que defende o pressuposto de um consumo televisivo mais ritualizado onde se observa uma forte influência da disponibilidade do público e dos factores de planeamento. Também, aborda a questão da “lealdade” aos suportes publicitários. A exposição encontra-se condicionada por múltiplos factores que podem contribuir para que as “preferências” não coincidam com o que é efectivamente maioritariamente seleccionado pelo público.

A “lealdade” poderá ser consequência de factores estruturais e/ou individuais. Nas bases de audimetria constam determinados indicadores que permitem avaliar quais são os suportes que se “associam” em maior grau com determinadas variáveis sociodemográficas, permitindo, assim, seleccionar os melhores suportes a incluir nos planos de *media*. Neste capítulo pretende-se a análises destes indicadores o que permitirá identificar quais são os suportes que apresentam uma maior adesão por parte dos diversos estratos sociodemográficos representados nas bases de audimetria. A selecção poderá não corresponder frequentemente a preferências e a multiplicidade de factores que condicionam a exposição televisiva contribui para uma maior irregularidade da exposição o que dificulta as análises de selecção dos suportes.

A caracterização do perfil de audiências de cada suporte condiciona o planeamento de *media*, uma vez que devem ser seleccionados os suportes que apresentem uma maior exposição ao público-alvo da campanha publicitária, sejam estes programas, tipologias ou canais. De acordo com *Fernandes (2001)*:

O perfil das audiências é um conjunto de informação relevante na definição do espaço publicitário, uma vez que os anunciantes têm públicos-alvo para os seus produtos. Consoante o perfil das audiências de cada canal e o cruzamento destes indicadores com os diferentes períodos horários, localizam-se os diferentes públicos.

Também, são necessárias abordagens metodológicas que o viabilizem e que são função das características dos dados, sendo descritos e implementados alguns métodos neste capítulo. No decorrer da última década, a eficácia da publicidade televisiva tem sido consideravelmente questionada, no entanto a televisão mantém-se como o *media* mais seleccionado para este fim. Com o intuito de avaliar o grau de adequação de um programa para a divulgação de determinado produto ou serviço será necessário considerar inúmeros factores, entre os quais o grau de envolvimento dos telespectadores com os programas, a lealdade da audiência para com determinados programas, canais e tipologias e a correlação existente entre os índices de exposição a determinado programa e as taxas de aquisição dos produtos ou serviços divulgados. No *Apêndice F.3* consta a análise do grau de envolvimento dos telespectadores e neste apêndice pretende-se a associação de suportes às variáveis sociodemográficas.

Na fase de optimização que integra a construção dos planos de *media* será necessário identificar o melhor conjunto de programas para promover determinada inserção publicitária de modo a maximizar as taxas de cobertura dos mesmos, atendendo às restrições orçamentais existentes. A selecção programática mais adequada, também, será função do grau de exposição dos telespectadores, da composição sociodemográfica do público e do grau de envolvimento com a programação. Alguns programas ou tipologias podem ser mais eficazes quanto se pretende captar e sustentar o interesse dos telespectadores, levando a menores distrações no decorrer dos seus espaços comerciais. As faixas horárias também podem diferir quanto ao grau de exposição dos telespectadores para com os espaços comerciais.

Os segmentos divergem quanto ao valor que as agências de publicidade lhes atribuem. Com base em variáveis sociodemográficas que caracterizam o público de determinados programas, as agências de planificação de meios comparam o plano elaborado com os consumidores alvo, por forma a avaliar se os objectivos definidos podem ser alcançados. Uma inserção publicitária será mais eficaz se a composição sociodemográfica da sua audiência apresentar uma elevada percentagem dos segmentos-alvo face à dimensão da audiência total a que serão expostos os programas seleccionados. Consequentemente, a análise da dimensão da exposição sociodemográfica aos canais, programas e tipologias e a análise dos comportamentos de lealdade e de fidelidade programática é determinante para um bom desempenho das agências de planificação de meios e de todos os agentes económicos que utilizam este *media* para divulgar os seus produtos ou serviços.

Na literatura internacional da especialidade encontramos alguns trabalhos que abordam a análise de alguns dos aspectos enumerados anteriormente, recorrendo em geral a metodologias de análise estatística tradicionais. Neste apêndice pretende-se avaliar as potencialidades de áreas metodológicas mais recentes, nomeadamente da Análise Simbólica e do *Data Mining*. As inserções publicitárias devem ser associadas aos programas e aos canais que são seleccionados com uma maior frequência por parte dos segmentos-alvo. As taxas de recordação das inserções publicitárias são função do grau de exposição dos telespectadores e da correspondente frequência ao longo de determinado horizonte temporal.

As metodologias simbólicas são as mais adequadas quando se pretende a análise da evolução temporal de determinados indicadores, assim como as metodologias de *Data Mining* que permitem a análise de representações dos dados com uma maior complexidade. Nomeadamente, os métodos que visam a análise de bases de dados transaccionais e que promovem a identificação do conjunto de programas seleccionados com uma maior frequência por parte dos consumidores-alvo. As análises sugeridas apresentam uma grande relevância prática para a elaboração dos planos de *media* e para a previsão dos índices de audiências. As estações televisivas vendendo espaços publicitários devem compreender qual é a dinâmica das escolhas realizadas, assim como as empresas que desenham as campanhas publicitárias dirigidas a determinados nichos de mercado.

A evolução observada nos principais indicadores de audimetria, também, se encontra dependente da composição sociodemográfica dos canais. Nomeadamente, quando um programa de grande popularidade é transmitido em determinado canal, os comportamentos de “lealdade” condicionam a capacidade dos canais em sustentar e atenuar os fluxos de audiência em direcção aos canais concorrentes, influenciando o grau de variabilidade observado nos principais indicadores que caracterizam o desempenho da grelha de programação.

Alguns dos trabalhos realizados aos índices de duplicação promovem a análise de comportamentos de lealdade dos telespectadores. Os académicos que pretendem perceber as motivações da audiência e os efeitos do consumo televisivo também têm procurado inter-relacionar o público e a grelha de programação. Os trabalhos que abordam as preferências e as escolhas programáticas têm-se debruçado nas motivações e nos gostos ou em características estruturais da programação. Em termos globais a abordagem estrutural tem sido mais frutífera, demonstrando que a audiência total e as variáveis temporais são os factores chave para prever as selecções realizadas. Isto sugere que todos os factores que influenciam a disponibilidade da audiência num dado instante são importantes para perceber as escolhas realizadas. Estes factores incluem além do período horário em que este é transmitido, quem está a assistir e qual o canal em que o programa é transmitido, entre outros.

Numa tentativa de justificar se quem assiste ao programa X também assiste ao programa Y, os resultados da maioria dos investigadores demonstraram que as melhores previsões não se baseiam na análise dos conteúdos programáticos. Com raras excepções a probabilidade de serem seleccionados em conjunto dois programas poderá ser prevista através dos respectivos *ratings*, sugerindo que o consumo é pouco mais do que o resultado de escolhas aleatórias.

No entanto, a investigação global realizada aos padrões de consumo verificaram que, por vezes, a audiência tende a ser leal a determinados canais ou programas e têm-se alcançado resultados contraditórios no que diz respeito à importância do conteúdo nas escolhas realizadas. Vários trabalhos têm concluído que o público que assiste a uma tipologia de programa não seleccionará necessariamente outro programa com as mesmas características, contrastando com a perspectiva de outros autores que defendem a existência de alguma lealdade às tipologias programáticas.

As tipologia programáticas são categorizações que permitem a produção continua de séries televisivas e fornecem heurísticas para a estimação do potencial sucesso dos programas propostos baseados nos resultados alcançados por programas dos mesmos géneros. As tipologias são um pilar da nomenclatura televisiva, uma vez que esta indústria se apoia fortemente num carácter repetitivo ao nível das produções. Neste contexto, os géneros serão um importante factor para a caracterização das escolhas dos telespectadores.

Alguns trabalhos suportam a existência de lealdade a tipologias, ou seja, defendem que o público que assiste a determinado programa apresentará uma maior tendência para seleccionar outros programas de idêntica tipologia. No entanto alguns autores caem no erro de confundir esta tendência com a lealdade programática. Os géneros televisivos diferenciam-se quanto ao possível grau de lealdade que despertam nos telespectadores. Alguns autores descobriram que os efeitos de herança entre programas adjacentes são superiores quando existem sinergias entre os géneros programáticos. Quando os programas não são transmitidos em faixas horária adjacentes é mais questionável, a importância das tipologias quando se pretende prever as taxas de audiências conjuntas.

Outros trabalhos defendem que excluindo as telenovelas e alguns programas diurnos, a audiência não apresenta lealdade a determinadas categorias, mas, sim, evita determinados conteúdos. Alguns autores são mais cautelosos a este nível, uma vez que a correlação existente entre os canais e os conteúdos programáticos favorece a possibilidade de se verificar a existência de lealdade programática. Na realidade não é fácil isolar o efeito decorrente das tipologias nas escolhas realizadas da influência dos restantes factores, no entanto é inquestionável que programas de determinadas tipologias são mais prováveis de ser vistos pelas mesmas pessoas.

Os trabalhos desenvolvidos no âmbito da lealdade aos canais pretendem isolar esta tendência, para além dos efeitos de herança e das selecções múltiplas existentes ao longo de uma sessão. Definem a lealdade a canais como a tendência do público para assistir consecutivamente a programas de determinado canal, em vez de distribuir o seu tempo por um conjunto de canais. A lealdade a um canal diminui com uma maior acessibilidade a um cada vez maior número de canais alternativos e, também, com a existência do controle remoto, que torna mais fácil a transição entre os canais concorrentes.

Os avanços tecnológicos na área dos *media*, sem dúvida que desincentivam à lealdade aos canais, no entanto existem evidências de que estes comportamentos ainda desempenham um papel fulcral na organização do consumo televisivo. Alguns trabalhos que têm tentado prever quais são as alterações observadas nas rotinas da exposição televisiva dos telespectadores, em função da disponibilidade e da acessibilidade a um cada vez maior número de canais, descobriram que o público tende a limitar as suas escolhas a um reduzido número deles, ou seja, desenvolvem lealdade a estes últimos. Será também necessário não esquecer que o comportamento individual não reflecte, apenas, as possíveis preferências individuais. Também, é função de outros factores, como, por exemplo, da exposição televisiva ao nível do agregado familiar, ou seja, das preferências dos restantes elementos que integram o grupo que rodeia determinado indivíduo. Os programas e os canais disponíveis no instante em que o telespectador decide ver televisão e o seu conhecimento a respeito de todas as suas opções, também, influencia os seus padrões de consumo.

Após a introdução realizada é possível concluir que a temática em análise é de grande complexidade, especialmente devido à dificuldade em isolar as distintas realidades que facilmente podem ser confundidas. Os dados que estão disponíveis são indicadores agregados e não medidas individuais o que não facilita e condiciona fortemente as análises a realizar e os resultados alcançados. A notoriedade de um canal condiciona os seus índices de audiência, uma vez que pode incentivar os telespectadores a seleccionar consecutivamente os seus programas, o que se denomina na literatura da especialidade de “lealdade” ao canal. No entanto, também não deixa de ser verdade que a estruturação da grelha e as tipologias transmitidas favorecem ou não esta predisposição. A dissociação dos efeitos decorrentes das exposições ou preferências por programas, tipologias e por canais é extremamente difícil, uma vez que interagem fortemente e condicionam-se reciprocamente.

F.5.2. Objectivos

Neste apêndice pretendemos a análise de alguns indicadores de audimetria que reflectem diferentes aspectos da exposição sociodemográfica ao meio televisivo. Um dos objectivos corresponde à associação de programas, tipologias e canais aos segmentos representados nas bases de dados disponibilizadas, o que permitirá a avaliação de comportamentos de “lealdade” e a definição do perfil das audiências dos suportes. Assim, também contribuímos para a selecção dos suportes a incluir nos planos de *media*, uma vez que devem apresentar uma elevada cobertura do público-alvo de uma campanha publicitária. Neste apêndice constam as análises realizadas a determinados índices sociodemográficos de atractividade, afinidade e adesão.

Através da associação de programas a segmentos pretendemos verificar qual é o conjunto de programas que é mais seleccionado e, assim, verificar se existe uma maior incidência em determinadas tipologias programáticas ou canais. A existência de comportamentos de lealdade permite a identificação dos programas mais adequados aos segmentos-alvo e avaliar aspectos relevantes para a previsão dos *ratings*.

Estes índices traduzem principalmente o grau de exposição da audiência a determinado programa, o que poderá não corresponder necessariamente a preferências concretas, mas, sim, à alternativa seleccionada atendendo às opções disponíveis em determinado momento.

No entanto, o que se pretende nas análises a realizar é a identificação dos programas seleccionados em maior grau por parte de cada um dos segmentos, ou seja, os programas que apresentam maiores índices de atractividade ou adesão e os programas que mais influenciam o público a permanecer no seu decorrer, o que se traduz em maiores valores de fidelidade ou de afinidade. Os índices de afinidade, uma vez que resultam do quociente de valores de *ratings*, ponderam cada telespectador quanto à duração do contacto e diferenciam-se dos índices de atractividade por dependerem dos níveis de audiências para o total do *media*, e, logo, estão condicionados pelos factores de disponibilidade do painel.

As exposições obviamente resultam da atitude dos telespectadores que será função de inúmeros factores, para além de eventuais preferências, como, por exemplo, de estratégias de programação e de contra-programação, de disponibilidade e necessidades, do planeamento publicitário, da notoriedade do canal e de hábitos adquiridos no decorrer do tempo. As preferências programáticas apenas podem ser avaliadas através da realização de inquéritos ao painel, não sendo o mais relevante para o objectivo principal proposto. Algumas conclusões nesta área são referidas por *Fernandes (2008)*.

Com base nas análises que foram sugeridas, talvez seja possível retirar algumas ilações pertinentes, ainda que condicionadas pelo facto dos índices em análise traduzirem medidas agregadas de quantificação da exposição televisiva. Não deve ser ignorado o facto destes indicadores traduzirem a exposição da audiência ao *media*, mas condicionada pelos diversos factores estruturais existentes. Logo, nem sempre traduzem preferências concretas, resultantes de um processo activo de selecção programática por forma a satisfazer determinadas necessidades individuais subjacentes. Por vezes, a audiência apresenta alguma passividade face à programação existente, sendo condicionada pelo seu conteúdo e pelas estruturas de planeamento. Pretende-se a identificação de possíveis tendências e de comportamentos de lealdade no consumo televisivo, através da associação de programas, tipologias e canais a determinados segmentos sociodemográficos.

Pretende-se, também, verificar se existem sequências programáticas que apresentem uma maior exposição a determinado conjunto de telespectadores, ajudando a prever quais serão as suas futuras selecções mais prováveis. Estas análises são especialmente relevantes para os segmentos que apresentem uma maior variabilidade nos principais indicadores de audimetria. As questões em avaliação correspondem às seguintes:

➤ Quais são os programas, tipologias programáticas e canais que podem ser associados em maior grau aos diferentes segmentos sociodemográficos incluído nas bases de dados de audimetria? Os canais diferem fortemente ao nível da composição sociodemográfica da sua audiência, nomeadamente os canais públicos face aos privados? Determinados segmentos apresentam uma maior exposição a determinadas tipologias ou existe alguma diversidade? Para realizar estas associações será preferível recorrer aos índices de atractividade, afinidade ou adesão? Quais são as abordagens metodológicas mais adequadas em função dos contextos, nomeadamente considerando programas com apenas uma transmissão ou com múltiplas transmissões? Quais são os suportes mais adequados para a promoção de determinados produtos, em função da sua exposição televisiva a determinados segmentos?

Quando se constroem planos de *media*, para além de ser necessário dispor de informação de caracterização do perfil das audiências dos suportes, também é necessário proceder à sua segmentação recorrendo aos dados primários de audimetria ou a dados agregados. Também, são sugeridas determinadas abordagens metodológicas que promovam a segmentação das audiências, comparando as potencialidades das metodologias de análise estatística e de *Data Mining*. Esta questão será desenvolvida no *Capítulo VI*.

F.5.3. Bases de dados

A base de dados corresponde aos índices de audiências que caracterizam a exposição do painel à grelha de programação de cada um dos segmentos sociodemográficos representados no painel de audimetria. Os índices mais relevantes para as análises a realizar são os seguintes: (1) Índices de Atractividade, (2) Índices de Adesão e os (3) Índices de Afinidade.

O índice de afinidade indica-nos se determinado suporte é idóneo para o público-objectivo pretendido. Este índice corresponde à percentagem da audiência sobre a audiência total do suporte que representa a audiência do público-objectivo pretendido. A afinidade permite-nos conhecer em que medida o nosso *target* se encontra exposto a esse suporte em comparação com a população total, indicando quais são os suportes que apresentam um menor desperdício de contactos relativamente ao público-objectivo da campanha.

O índice de atractividade desempenha a mesma função, mas, não, se encontra influenciado pelos níveis da exposição global ao *media*, uma vez que recorre a valores de *shares* de audiências. Os *shares* representam o *ranking* das audiências dos programas e das diversas faixas horárias, de forma similar a uma quota de mercado. No seguinte *quadro* consta a caracterização dos indicadores enumerados anteriormente (*Quadro F.5.3.1*):

Indicador	Descrição	Fórmula de Cálculo
Índice de Atractividade	Compara o <i>Share</i> de um programa para um segmento com o valor do <i>Share</i> obtido por esse mesmo programa para o segmento de referência (em geral para o universo).	$Att = \frac{Shr\_Prog\_X\_Seg\_Y}{Shr\_Prog\_X\_Univ} * 100$
Índice de Adesão	Mede a estrutura da audiência em um evento, ou seja, a proporção de indivíduos de determinado <i>target</i> na audiência total do evento.	$Adh = \frac{Rat\_Prog\_Seg\_Y * N^o of Ind\_Seg\_Y}{Rat\_Prog\_Univ * N^o of Ind\_Univ} * 100$
Índice de Afinidade	Diferencia-se do índice de atractividade por se referir a valores de <i>Ratings</i> em vez de a valores de <i>Shares</i> .	$Aff = \frac{Rat\_Prog\_Seg\_X}{Rat\_Prog\_Univ} * 100$

Quadro F.5.3.1 – Caracterização das variáveis que traduzem a exposição sociodemográfica e que viabilizam a selecção dos melhores suportes a incluir nos planos de *media*

As unidades estatísticas nas análises a realizar correspondem a programas, considerando o primeiro nível programático ou a períodos horários caracterizados pelos índices anteriormente referidos para cada um dos segmentos representados nas bases de dados. O painel encontra-se segmentado em função das variáveis: classes sociais, regiões, idades, sexo, lares com ou sem TV Cabo e ser Dona de Casa, isto é, responsável pelas compras do lar (ver Apêndice A para descrição das bases de dados). Considera-se a grelha de programação dos canais generalistas. Nas bases de dados simbólicas, os conceitos correspondem a faixas horárias com trinta minutos de duração ou a um conjunto de episódios de determinado programa caracterizados por variáveis essencialmente da tipologia intervalo.

F.5.4. Metodologias

(I) Abordagens metodológicas utilizadas para associar segmentos sociodemográficos a suportes publicitários

As metodologias a utilizar procedem da Estatística Multivariada, do *Data Mining* e da Análise Simbólica. A maioria das análises correspondem a questões que se traduzem no *clustering*, na classificação de observações e na redução de dimensionalidade dos dados e respectiva representação. Algumas das metodologias utilizadas foram as que constam do *Quadro.F.5.4.1*.

Para a análise dos índices de atractividade as abordagens metodológicas diferenciam-se em função da tipologia da amostra de programas. Para os programas com uma única transmissão, por exemplo, filmes, jogos de futebol e galas adoptaram-se duas abordagens: O MDS→RLM e o Clust→ Discriminante e para a análise de dados simbólicos referentes a programas com mais de um episódio recorreu-se à análise ACP simbólica.

No método MDS→RLM utiliza-se a matriz de dissemelhança entre programas para realizar uma análise MDS não métrica, com a posterior extracção das variáveis mais relevantes para a quantificação das dissemelhança por aplicação da Regressão linear e finaliza-se com uma análise de agrupamento sobre essas variáveis.

Metodologias de Estatística Multivariada	Metodologias de Análise Simbólica (Bock & Diday, 2000)
Análise MDS, a análise Anova (Hinkelmann, Klaus & Kempthorne, 2008) ou Kruskal Wallis (Corder & Foreman, 2009) no caso não paramétrico, Teste T para amostras independentes ou Mann - Whitney U no caso não paramétrico, a Regressão Linear Múltipla, a Análise Discriminante Linear (Reis, 1997), o algoritmo de <i>clustering</i> de K Média (Ding & Xiaofeng, 2004) e a análise de <i>Clusters</i> Hierárquica (Reis, 1997).	As <i>Zoom Stars</i> , as Redes Neurais de Kohonen para dados simbólicos, as medidas de Dissemelhança Simbólicas para objectos simbólicos booleanos e a análise em Componentes Principais.
Metodologias de Data Mining	
Alguns algoritmos de <i>clustering</i> , nomeadamente o K Média e algoritmos de associação, como, por exemplo, o Apriori.	

Quadro F.5.4.1 – Metodologias utilizadas nas análises de associação de suportes a segmentos sociodemográficos

Vão-se construir mapas perceptuais com base na análise MDS não métrica baseados nos índices de atractividade dos programas para determinados segmentos sociodemográficos. Os programas que são mais seleccionados por determinado segmento localizar-se-ão próximos no espaço MDS calculado e os programas que se diferenciem quanto ao seu público afastar-se-ão. A densidade programática poderá ser mais elevada em algumas regiões, o que pode indiciar a percepção de semelhanças por parte do seu público. Quando programas aparentemente dissimilares encontram-se localizados de forma próxima, isto poderá ser atribuído à complementaridade procurada pelo seu público. Associando-se programas às variáveis de segmentação do painel, talvez se possa concluir também a respeito das tipologias programáticas e dos canais que foram mais seleccionados.

A utilização da Regressão Linear Múltipla irá ajudar na interpretação das dimensões dos mapas perceptuais, ou seja, permite a extracção de variáveis relevantes para a quantificação das dissimelhanças. Estas variáveis podem ser representadas no espaço das dimensões MDS, utilizando os coeficientes de regressão padronizados. Seria desejável observar uma elevada dispersão das variáveis no espaço perceptual de forma às variáveis identificadas representarem fielmente as dimensões MDS.

No método CLUST → DISCRIMINANTE utilizamos directamente a matriz de dissimelhança para definir os agrupamentos. Posteriormente recorre-se a uma metodologia de análise Discriminante Linear ou não linear para a selecção das variáveis utilizadas na caracterização e interpretação dos grupos anteriormente definidos.

A análise Discriminante Linear ou o algoritmo CART de classificação conjuntamente com a análise de *Cluster* e o K Média constitui uma alternativa aos mapas perceptuais nas situações em que se verificou uma maior dificuldade na sua interpretação. A avaliação do poder discriminante global das variáveis de classificação corresponde na análise Discriminante Linear ao cálculo do índice de potência representado por

$$PI(X_j) = \frac{\sum_{i=1}^k \hat{a}_{ji}^2 \lambda_i}{\sum \lambda_i}$$

No seu cálculo utiliza a contribuição de cada variável para a função discriminante e a contribuição da função para a solução global. Corresponde ao somatório do produto do quadrado dos coeficientes estruturais e do λ de cada função discriminante. K será o número de funções discriminantes e $\hat{a}_{ji}$ corresponde ao coeficiente estrutural da variável X_j para a função i. Nas situações em que se opte por recorrer às Árvores CART a medida de relevância global das variáveis corresponde a

$$M(X_j) = \sum_{O \in A} z^{j0} \Delta I(\Pi^j, O, A) \quad \text{com} \quad z^{j0} = \begin{cases} 1, & \Pi^j \in \Pi^s(O) \\ 0, & \Pi^j \notin \Pi^s(O) \end{cases}$$

onde, Π^j representa a ramificação do nó O associada com a variável de classificação X_j , Π^s representa o conjunto de variáveis substitutas e $\Delta I(\Pi^j, O, A)$ traduz o decréscimo numa medida de impureza na Árvore global (A) conseguida com a utilização da variável X_j quando se procede à partição das observações no nó O. A medida estandardizada corresponde a

$$M(X_j) = 100 * \left[\frac{M(X_j)}{\max M(X_j)} \right]$$

No método ACP simbólico em primeiro lugar constrói-se uma matriz de dissemelhanças simbólica com base nos intervalos de variabilidade dos índices de atractividade programáticos para cada um dos segmentos. Posteriormente, aplica-se a análise em Componentes Principais Simbólica para variáveis da tipologia intervalo.

Para programas com mais de uma transmissão ao longo do período em análise, por exemplo, espaços informativos, telenovelas, séries de ficção e *talk-shows*, uma vez que não dispomos de uma única observação por programa faz todo o sentido recorrer a metodologias de Análise Simbólica. Os programas que forem transmitidos em diferentes períodos horários ou simultaneamente ao longo da semana e ao fim-de-semana agruparam-se segundo o período correspondente. Os objectos de segunda ordem correspondem, portanto a subconjuntos de programas caracterizados por variáveis da tipologia intervalo, o que corresponde a objectos simbólicos booleanos.

Nas análises também vão ser utilizadas as Zoom Stars a 2 e 3 dimensões, que permitem visualizar os intervalos de variação dos índices de atractividade para os vários segmentos sociodemográficos. Também, se irá utilizar a análise em Componentes Principais Simbólica para dados da tipologia intervalo em determinadas situações. Este algoritmo produz resultados semelhantes ao Método dos Vértices da análise PCA simbólica. Espera-se que os factores estejam essencialmente associados ao canal, dias e períodos horários. As correlações não devem ser exclusivamente atribuídas ao conteúdo do programa.

Por conseguinte, os modelos que têm sido desenvolvidos no sentido de associar aspectos sociodemográficos a tipologias não têm produzidos bons resultados quando utilizados por si só para prever as escolhas realizadas. Para a análise dos índices de adesão e de fidelidade vamos utilizar as Redes Neurais SOM simbólicas, assim como as Regras de Associação e as Sequenciais.

Com base nos algoritmos de *clustering* pretendemos identificar agrupamentos de programas com base nas semelhanças observadas para estes indicadores, os quais se encontram caracterizados em função dos seus protótipos. Uma vez que o objectivo das análises não corresponde à previsão de índices de audiências para novos programas, não é necessário a identificação de variáveis que permitam afectar novos programas a uma das tipificações pré definidas. Se assim fosse seria necessário recorrer à Análise Discriminante.

As Regras de Associação permitem verificar quais são os conjuntos de programas seleccionados com uma maior frequência por parte de determinados segmentos. Recorreu-se a esta metodologia nas situações em que as metodologias de *clustering* não foram eficazes. As análises que visam a utilização de Regras de Associação deveriam incidir em bases de dados transaccionais, como o são as bases de *raw data* de audimetria. Essas análises permitem prever as selecções programáticas com base nos dados históricos referentes à sequência de programas que foram seleccionados. A cada elemento do painel associamos a data e hora em que selecciona determinado programa. Uma vez que esta tipologia de dados não foi disponibilizada, podemos considerar alternativamente dados agregados de segmentos sociodemográficos.

Os eventos vão corresponder a segmentos e aos programas que apresentam valores de afinidade superiores à unidade ou no 4º quantil no decorrer de determinado dia, quer dizer, a <Segmento_Id, Data, Horizonte Temporal_Y, Canal_X, {Prog_X, Prog_Y, ... , conjunto de programas para os quais o segmento identificado apresenta valores de adesão superiores à unidade}>. Também, podemos utilizar variáveis relativas aos canais, como, por exemplo, <Canal_X, Data, Horizonte Temporal, {Tempo_Total_Despendido no Prog_X ou na Tipologia_Programática_Y; Ratings para Prog_X ou Tipologia_Y}>.

Recorrendo ao algoritmo Apriori do *software* SAS derivamos Regras sequenciais na qual a ordem dos programas será relevante. Como resultado do algoritmo obtemos as Regras com maiores valores de frequência e confiança. $A \rightarrow B$ será uma sequência de programas que maximizam o indicador em análise para os segmentos sociodemográficos seleccionados e $N_{A \rightarrow B}$ corresponde ao número de dias em que essa sequência surge e N será o número total de dias. O *Link Analysis* permite visualizar as relações existentes e a identificação de quais são os programas que partilham um maior número de telespectadores e a análise MDS poderá, uma vez mais, reproduzir as dissemelhanças existentes num espaço eucladiano bidimensional.

Os índices de afinidade representam a evolução dos *ratings* sociodemográficos comparativamente aos associados à totalidade do painel, permitindo a identificação de programas e de períodos onde se destacam face ao global. Em primeiro lugar associam-se intervalos desta variável a cada um dos programas com múltiplas transmissões e aplicam-se metodologias simbólicas, como, por exemplo, a análise ACP. Posteriormente, para as análises que visam programas de um único episódio recorrem-se a algoritmos de *clustering* de *Data Mining*.

(II) Abordagens metodológicas de segmentação das audiências televisivas e de selecção dos suportes

Em Kim (2002) consta o seguinte processo de segmentação das audiências televisivas: ETAPA I – Recolha de dados dos ficheiros base de Audimetria; ETAPA II – Cálculo de um índice de correlação que avalia a probabilidade de determinado par de programas ser seleccionado por um telespectador. Aplicação da análise de *clusters* para identificar sub-mercados de programas; ETAPA III – Aplicação do K Média para identificar os segmentos de telespectadores utilizando como semente os segmentos de programas identificados anteriormente e análise do perfil sociodemográfico de cada segmento identificado.

Como alternativas metodológicas referem-se as seguintes opções:

Segmentação de audiências recorrendo aos dados dos ficheiros de base de audimetria	Segmentação recorrendo a dados agregados de audimetria
ETAPA I – <i>Clustering</i> de registos de exposição televisiva, identificando os registos semelhantes, por aplicação de algoritmos de <i>Clustering</i> de bases de dados transaccionais. Implicitamente, também se podem identificar algumas associações entre os suportes. ETAPA II – Derivar Regras de Associação ou Sequenciais, verificando para cada segmento previamente identificado quais são as associações mais características entre os programas. Também podemos trabalhar com faixas horárias ou com tipologias programáticas. Para além, de identificar quais são os programas vistos com maior frequência nas sessões televisivas dos telespectadores, também podemos identificar quais são os programas que implicam a ausência de selecção de outras tipologias programáticas. Posteriormente, a caracterização de cada segmento recorrendo a variáveis sociodemográficas, como, por exemplo, por aplicação de metodologias simbólicas.	Segmentação de audiências recorrendo a dados agregados, ou seja, a indicadores de audimetria, mas, também, podemos recorrer a dados de fluxos de audiências por segmentos Podem ser seleccionados para macro-segmentos os programas que apresentam índices de atractividade ou de adesão superiores a um valor a definir e, assim, identificando os programas que apresentam maiores associações no decorrer de um dia. Nestes contextos as abordagens metodológicas anteriormente definidas são válidas. Alternativamente, as associações tipológicas e programáticas têm que ser realizadas através da construção de mapas perceptuais recorrendo a metodologias clássicas ou simbólicas.

Quadro F.5.4.2 – Metodologias que contribuem para a segmentação das audiências televisivas

No capítulo VI desenvolve-se esta temática em especial na avaliação da quarta questão onde se refere as abordagens alternativas para a construção dos planos de *media*.

F.5.5. Resultados

I. Análise aos índices de atractividade

Neste capítulo constam as análises realizadas aos índices de atractividade, os quais comparam o *share* de um evento para determinado segmento com o valor do *share* para o alvo base, tendo-se considerado o universo. Procurou-se associar programas, tipologias e canais aos diversos segmentos representados nas bases de dados de audimetria e exemplificar métodos de selecção de suportes quando se constroem os planos de *media*.

Os resultados diferenciam-se em função das características da programação, nomeadamente entre programas de uma única transmissão vs. programas transmitidos em episódios. Para programas com mais de um episódio recorreremos a metodologias simbólicas, como, por exemplo, à análise ACP e às Zoom Stars uma vez que agregamos os valores da variável em estudo.

Para os restantes programas realizam-se análises MDS, recorrendo à RLM para uma melhor interpretação das dimensões subjacentes ou à análise de *Clusters*, K Média e à análise Discriminante Linear. As unidades estatísticas correspondem a programas caracterizados pelos índices de atractividade.

Os resultados diferenciam-se em função do canal considerado e da tipologia de programas, ou seja, de uma única transmissão versus múltiplas transmissões.

Ia. Resultados para o canal TVI e para programas com múltiplos episódios

Realizando uma análise em Componentes Principais simbólica, segundo o método dos vértices, considerando a globalidade das variáveis e de objectos simbólicos será necessário reter cinco eixos para obter uma percentagem acumulada de inércia explicada de 77%. Ao primeiro factor corresponde 35,4% e ao segundo 19%.

Não constitui como um bom conjunto de variáveis para diferenciar os vários objectos simbólicos, que corresponde ao círculo de correlações das variáveis com os eixos factoriais.

Os dois principais eixos opõe dois conjuntos de variáveis: as faixas etárias, donas de casa, lares com TV cabo e um segundo conjunto englobando regiões, sexo e as classes sociais.

➤ 1º Factor

As variáveis menos correlacionadas com o primeiro factor são os segmentos Costa Norte, o Interior, a classe D, as mulheres, as faixas etárias dos 4-14, 45-54, 55-64, +64, Lares sem TV Cabo e as donas de casa (*Gráfico F.5.5.I.1*).

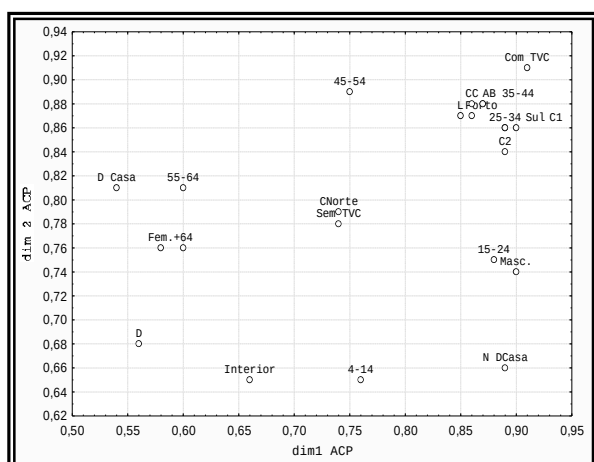


Gráfico F.5.5.I.1 - Correlações das variáveis com os eixos para a análise ACP Simbólica realizada com os programas da TVI (idêntico à ACP clássica, algoritmo dos vértices)

Portanto, podemos associar ao primeiro factor essencialmente as classes sociais mais elevadas, os adultos mais jovens até aos 45 anos, o sexo masculino, os lares que têm TV cabo, os telespectadores que não são donas de casa e todas as regiões com excepção do Interior e da Costa Norte.

Os objectos simbólicos melhor representados pelo primeiro eixo incluem a maior parte dos espaços destinados à transmissão de filmes ao fim-de-semana e também dos dias laborais transmitidos no período da noite, assim como as telenovelas e as séries da noite e do horário nobre. Inclui um maior número de programas transmitidos ao fim-de-semana.

➤ 2º Factor

O segundo factor está menos correlacionado com as regiões da Costa Norte, o Interior, a classe social D, ambos os sexos, a faixa etária dos 4-24 anos e a partir dos 55 anos, lares sem TV cabo e ser dona de casa.

Ao segundo factor podemos associar as mesmas regiões que foram identificadas para o anterior factor, as classes sociais mais elevadas, as faixas etárias entre os 24 anos e os 55 anos e os lares subscritores de TV Cabo. As variáveis sexo e a situação no lar não influenciam estes resultados. Diferencia-se do primeiro factor por não ser influenciado pela variável sexo e por ser dona de casa e está em maior grau associado aos telespectadores que apresentam idades compreendidas entre os 45 e os 55 anos. Os objectos simbólicos melhor representados pelo segundo eixo factorial correspondem a programas matinais, programas de pré-prime-time e apenas a um espaço destinado à transmissão de filmes ao Domingo à noite. Inclui programas que serão mais dificilmente acessíveis a quem exerce uma actividade profissional.

➤ 3º Factor

O terceiro factor está menos correlacionado com Lisboa, Costa Centro, Interior, Sul, classe C1 e classe D, ambos os sexos, faixas etárias dos 4 aos 24 anos ou com mais de 55 anos, lares sem TV Cabo e com a situação no lar. Portanto, está mais associado ao Norte, à classe A/B e C2, a adultos entre os 25 e os 55 anos e aos lares com TV cabo. Os objectos simbólicos melhor representados correspondem aos filmes transmitidos em alguns dos dias laborais e à telenovela que antecede o espaço informativo.

➤ 4º Factor

O quarto factor encontra-se correlacionado com as regiões do Porto e do Sul, todas as classes sociais excluindo a D, o sexo masculino, idades compreendidas entre os 25 e os 54 anos e com os lares com e sem TV cabo.

Portanto, o espaço informativo da noite e o espaço informativo da tarde ao fim-de-semana associa-se às classes sociais mais elevadas, a todas as regiões, excluindo o Interior, e aos adultos com idades compreendidas entre os 14 e os 55 anos. O espaço informativo da tarde em dias laborais está melhor representado pelo quarto factor e o espaço equivalente de fim-de-semana pelo 12º factor. O espaço informativo da noite da semana e de fim-de-semana está melhor representado pelo 12º factor que se encontra menos correlacionado com o Interior, a classe D, ambos os sexos, idades 4-14 e a partir dos 55 anos e com a situação no lar. O espaço informativo da tarde da semana associa-se ao Porto e ao Sul, a todas as classes excluindo a D, ao sexo masculino, faixas etárias dos 25 aos 54 anos, lares com ou sem TV cabo.

Ib. Programas da TVI transmitidos num único episódio

Para os programas com um único episódio construíram-se mapas perceptuais baseados no algoritmo MDS não métrico. Como matriz de entrada utilizou-se a matriz de correlações, que traduz as (di)semelhanças existentes entre os diferentes programas caracterizados pelos índices de atractividade para os vários segmentos. Diferenciam-se os resultados para determinados segmentos sociodemográficos: classes sociais, faixas etárias, sexo e situação no lar.

➤ As Classes Sociais

Para um modelo a três dimensões obteve-se um valor do Stress com o valor de 0,0000048 e para duas dimensões de 0,189811.

Os mapas perceptuais (*Gráfico F.5.5.I.2 e F.5.5.I.3*) baseiam-se nos índices de atractividade das classes sociais associados a uma amostra de programas transmitidos num só episódio, como, por exemplo, os filmes transmitidos ao longo de toda a semana em diferentes períodos horários e de diversas tipologias programáticas ou os jogos de futebol, entre outras tipologias.

No *Quadro F.5.5.I.1* estão identificadas algumas observações. Os programas a azul correspondem a maiores exposições por parte da classe social A/B, a lilás constam os programas associados à classe C1, a verde os associados à classe C2 e a rosa os associados à classe D.

Para auxiliar a interpretação das dimensões dos mapas perceptuais efectuaram-se análises de Regressão Linear Múltipla baseadas na matriz original dos dados que cruza programas e índices conjuntamente com as coordenadas de cada programa em cada uma das dimensões.

A variável dependente em cada uma das análises corresponde a cada um dos índices de atractividade e as variáveis independentes às coordenadas dos diversos programas para as várias dimensões. Os segmentos que estão mais relacionados com a solução proporcionada pelo MDS, estão associados a regressões com elevados coeficientes de determinação e devem apresentar uma ponderação mais elevada em algumas das dimensões MDS (*Tabela F.5.5.I.1*).

Com base na interpretação dos mapas perceptuais podemos fazer as seguintes associações: (1) Classe A/B e C2 apresentam uma maior exposição a filmes do fim-de-semana; (2) Classe C1 apresenta uma maior exposição a filmes ao longo da semana; (3) a Classe D apresenta uma maior exposição a filmes transmitidos à Sexta e ao fim-de-semana.

No que diz respeito às tipologias programáticas, para nenhuma das classes sociais existe aparentemente uma maior exposição a determinadas categorias.

Realizando uma análise Anova a dois factores em que se consideram as tipologias e as classes sociais como factores para uma amostra completa e com desigual número de observações por célula e considerando os mesmos programas das análises anteriores constatamos que apenas o efeito principal associado à tipologia de grau 3 é significativo (p valor de 0,0000). O efeito classe social apresenta um p valor de 0,09 e o efeito de interacção um p valor de 0,99.

Portanto, podemos concluir que o factor mais importante é na realidade a disponibilidade da audiência, não existindo uma clara associação de tipologias às classes sociais. A selecção programática não parece ser em maior grau função de “preferências” específicas por programas ou por tipologias programáticas.

Segundo o teste de Bartlett apenas para o efeito principal da tipologia se rejeita a hipótese nula com um p valor de 0,002.

Não existe correlação entre as médias e as variâncias. Em relação à normalidade dos grupos definidos pela combinação dos níveis associados aos dois factores em análise existem alguns desvios por observação dos histogramas e do gráfico normal. Para os restantes grupos em virtude da dimensão das amostras pelo TLC assume-se a normalidade da distribuição da variável dependente.

Segundo o teste HSD para dimensões de amostras desiguais verificamos que a tipologia de Fantasia distingue-se de todas as restantes tipologias com excepção do histórico. Estas duas tipologias maximizam os valores de *ratings*, após os jogos de futebol e correspondem a dois filmes de fim-de-semana de grande audiência. A tipologia de Aventura apresenta um menor índice médio de audiência face à Fantasia ou ao futebol e supera os filmes de Terror e de Ficção Científica. As tipologias associadas ao maior número de filmes, ou seja, drama, comédia, apresentam diferenças estatisticamente significativas para as classes sociais.

Realizando análises Anova e cruzando o factor classe social com os factores dia, semana/fim_semana e período horário apenas os efeitos principais associados aos factores temporais são significativos para um nível de significância de 5%. No caso da variável período, com os correspondentes níveis: Noite (19-22h) ($\mu=9,3$), Tarde (13-14.30h) ($\mu=4,6$), Final de Noite (22-24h) ($\mu=1,6$) e Final de Tarde (14.30-19h) ($\mu=6,8$) segundo o Teste de Scheffe apenas o par constituído pelo final da tarde e pela noite não se diferenciam entre si (p valor de 0,42).

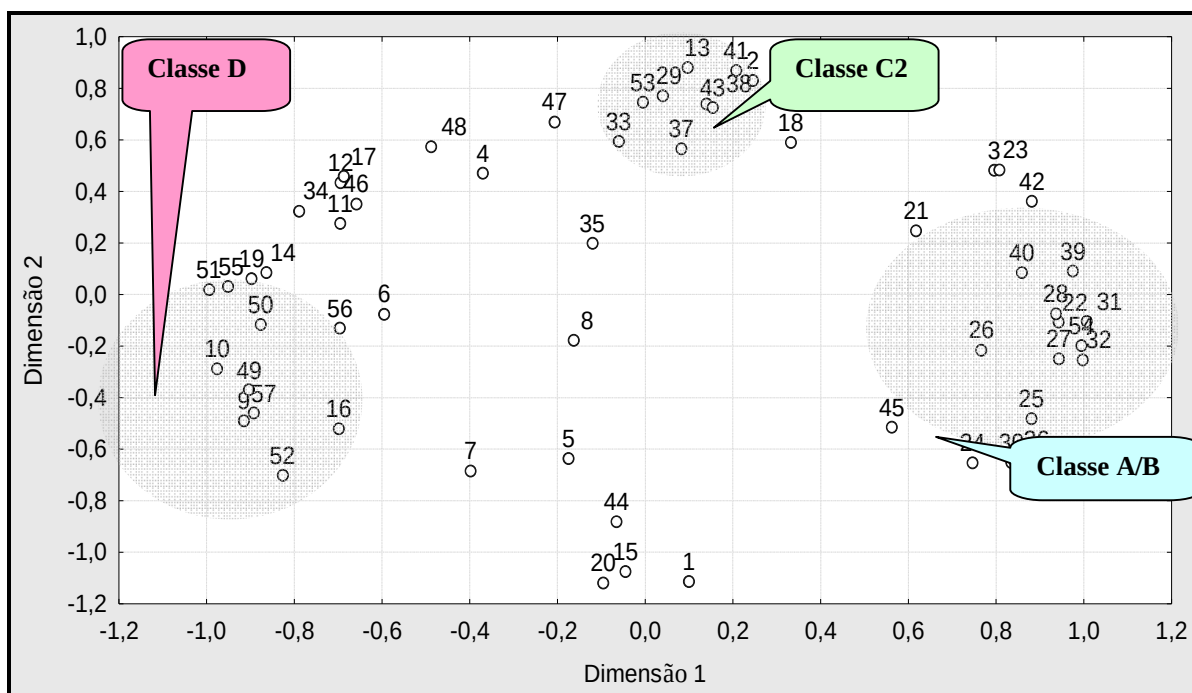


Gráfico F.5.5.I.2 - Mapa perceptual para as classes sociais e programas do Canal TVI (Dim. 1 e 2)

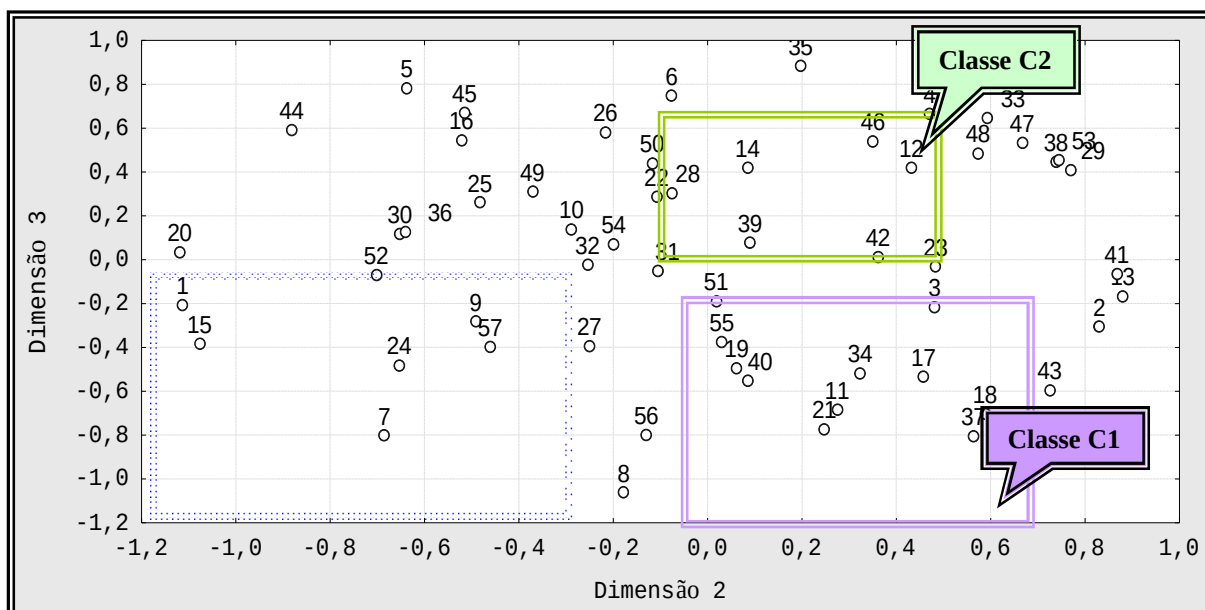


Gráfico F.5.5.I.3 - Mapa perceptual para as classes sociais e programas do canal TVI (Dim. 2 e 3)

Segmento	R Múltiplo	R ²	P Valor	Dim 1	Dim 2	Dim 3
A/B	0,86	0,66	0,0	0,74	-0,35	
C1	0,90	0,70	0,0		0,58	-0,61
C2	0,13	0,73	0,0	0,29	0,30	0,75
D	0,76	0,70	0,0	-0,66	-0,50	-0,15

Tabela F.5.5.I.1 - Interpretação das dimensões dos mapas perceptuais

Consequentemente, os períodos podem-se ordenar segundo a seguinte ordem crescente: final de noite, tarde e final de tarde com noite. O termo de interacção com a classe social apresenta um p valor de 0,98. Para o factor dia apenas existem diferenças entre o Sábado ($\mu=4$) e o Domingo ($\mu=6$) e entre eles e os restantes dias que apresentam uma média muito inferior ($\mu\approx 1,5$). Para o factor S/FS o período ao longo da semana (S) tem uma média de 1,7% e o fim-de-semana (FS) uma média de 5%.

Para este período e considerando esta tipologia de programas do canal TVI, os *ratings* das classes sociais são superiores para o fim-de-semana (o Domingo supera o Sábado), noite e final de tarde e jogos de futebol. As classes sociais não se diferenciam e não existe interacção com os restantes factores.

N.º	PROGRAMA
1	FILME DE SEGUNDA \ "DUELO IMORTAL"
2	FILME DE SEGUNDA \ "GOLPE QUASE PERFEITO"
3	FILME DE SEGUNDA \ "OS INSPECTORES "
4	FILME DE SEGUNDA \ "PAIXÃO SELVAGEM"
5	FILME DE SEGUNDA \ "ROCKY V"
6	FILME DE SEGUNDA \ "THELMA E LOUISE"
7	FILME DE SEXTA \ "AMEAÇA EXTRATERRESTRE"
8	FILME DE SEXTA \ "O DESPERTAR DO PASSADO"
9	FILME DE SEXTA \ "ESQUECER PARIS"
10	FILME DE SEXTA \ "STAR TREK – GERAÇÕES"
11	FILME DE TERÇA \ "CULPA FORMADA"
12	FILME DE TERÇA \ "LIGAÇÃO YAKUZA"
13	FILME DE TERÇA \ "OPERAÇÃO DE RISCO"
14	FILME DE TERÇA \ "ROCKY IV"
15	FILME DE TERÇA \ "SILÊNCIO DE MORTE"
16	FILME DE TERÇA \ "UM PEIXE CHAMADO VANDA"
17	FILME DE QUARTA \ "SEM MEDO DE VIVER"
18	FILME DE QUARTA \ "SEXO, MENTIRA E OBSESSÃO"
19	FILME DE QUARTA \ "TEMPO DE DECIDIR"
20	FILME DE QUARTA \ "VITIMAS DA JUSTIÇA"
21	FILME DE QUINTA \ "ANGIE - UMA MULHER SÓ"
22	FILME DE QUINTA \ "RISCO CALCULADO"
23	FILME ESPECIAL DE SÁBADO \ "HARRY POTTER"
24	MATINE DOMINGO I \ "007 - ALVO EM MOVIMENTO"
25	MATINE DOMINGO I \ "BRAVEHEART"
26	MATINE DOMINGO I \ "CIDADE EM PERIGRO"
27	MATINE DOMINGO I \ "HELENA DE TROIA"
28	MATINE DOMINGO I \ "MISSÃO IMPOSSIVEL"
29	MATINE DOMINGO II \ "GODZILLA"
30	MATINE DOMINGO II \ "JUSTIÇA VERMELHA"
31	MATINE DOMINGO II \ "O CACADOR DE TUBARÕES"
32	MATINE DOMINGO II \ "PERFUME DE MULHER"
33	MATINE DOMINGO III \ "LIMITE VERTICAL"
34	MATINE SABADO I \ "007 - MISSÃO ULTRASECRETA"
35	MATINE SABADO I \ "COMO AGARRAR UMMARIDO"
36	MATINE SABADO I \ "O REGRESSO DO PESTINHA"
37	MATINE SABADO I \ "OPERAÇÃO FLECHA"
38	MATINE SABADO II \ "007 - GOLDENEYE"
39	MATINE SABADO II \ "007 – O AMANHÃ NUNCA MORRE"
40	
41	MATINE SABADO II \ "A MUMIA"
42	MATINE SABADO II \ "O REGRESSO DA MUMIA"
43	MATINE SABADO II \ "ROMANCE PERIGOSO"
44	
45	FILME do DOMINGO \ "ALVO DE TRAIÇÃO"
46	FILME do DOMINGO \ "CÓDIGO VERMELHO"
47	FILME do DOMINGO \ "MEDO DE MORTE"
48	FILME do DOMINGO \ "RISCO CALCULADO"
49	FILME do DOMINGO \ "UMA VIAGEM SEM FIM"
50	
51	FILME do SABADO \ "CAÇA AO HOMEM"
52	FILME do SABADO \ "FENÓMENO"
53	FILME do SABADO \ "PLAYERS CLUB"
54	FINAL TAÇA DE PORTUGAL
55	FUTEBOL - TAÇA PORTUGAL \ "BENFICA-FC PORTO"
56	LISBOA REGRESSA AO PARQUE (R)
57	PROCISSÃO DAS VELAS
58	VIVA O ZOO

Quadro F.5.5.I.1 – Identificação de alguns números que constam nos mapas perceptuais

➤ Faixas Etárias

Segue-se uma idêntica análise considerando como variável de segmentação as faixas etárias. Na *Tabela F.5.5.I.2* e no *gráfico F.5.5.I.4* constam os resultados das análises de regressão realizadas para facilitar a interpretação de cada uma das dimensões MDS. No *Quadro F.5.5.I.2* consta a identificação dos números que representam os programas na solução MDS proporcionada. A verde estão os programas associados à faixa etária dos 35/44, a amarelo dos 45/54 anos, a azul 55/+64 anos, a rosa 4/14 anos, a lilás 15/24 anos e a vermelho 25/34 anos.

Para os espaços em que se transmite filmes ao fim-de-semana à tarde prevalece a exposição às faixas etárias dos 15 aos 44 anos e dos 55 aos 64 anos.

No que diz respeito aos filmes da noite não existe uma maior exposição por parte de determinados segmentos. Ao longo da semana estes programas exercem uma maior atractividade sobre as faixas etárias dos 15 aos 24 anos e às faixas etárias dos 35 aos 44 anos. Em relação às tipologias não existe uma clara associação com as faixas etárias.

➤ TV Cabo

Atendendo à variável que traduz o acesso à TV Cabo, curiosamente os índices de atractividade para os que lares que não têm TV Cabo são mais reduzidos para a generalidade dos filmes. A fracção do público que dispõe de TV Cabo apresenta maiores índices de atractividade para filmes de acção ao longo de toda a semana.

Segmento	R Múltiplo	R ²	P Valor	Dim 1	Dim 2	Dim 3
4/14	0,82	0,68	0,00	-0,69	-0,44	
15/24	0,52	0,23	0,00			0,52
25/34	0,74	0,52	0,00	-0,29	0,67	
35/44	0,77	0,57	0,00		0,37	-0,66
45/54	0,48	0,23	0,00			-0,45
55/64	0,62	0,39	0,00	0,59	-0,22	
+64	0,74	0,56	0,00	0,42	-0,60	

Tabela F.5.5.I.2 - Interpretação das dimensões para a análise MDS que associa programas a faixas etárias

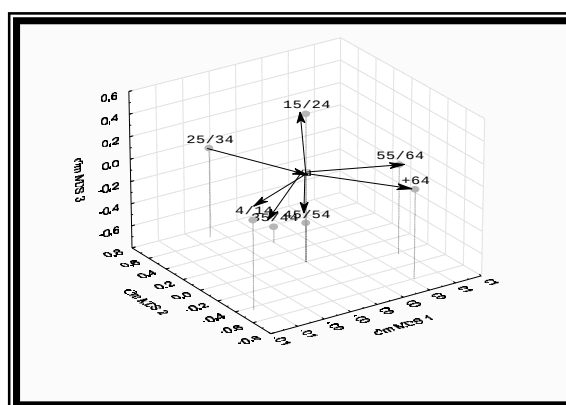


Gráfico F.5.5.I.4 – Variáveis estatisticamente significativas na regressão com as dimensões MDS

Com o objectivo de associar programas aos segmentos definidos com base na variável Sexo e responsável pelas compras do agregado familiar, em substituição das análises MDS que não permitiram retirar conclusões pertinentes, aplicou-se a análise de *Cluster* para agregar os programas em função das dissimilaridades apresentadas e, posteriormente recorreu-se à análise Discriminante Linear para identificar os segmentos que permitem diferenciar esses mesmos grupos. As observações estão identificadas no *Quadro F.5.5.I.3* e correspondem à amostra utilizada nas análises anteriores.

A rosa constam os programas associados aos telespectadores que não são donas de casa, a vermelho aos telespectadores que são donas de casa, a azul aos do sexo masculino e não são donas de casa e a preto ao sexo feminino e a donas de casa. A matriz de entrada utilizada para a análise de *clusters* foi a matriz de correlações entre os programas caracterizados em função dos seus índices de atractividade para os segmentos enumerados anteriormente. Uma vez que as variáveis estão medidas na mesma escala não é necessário a sua standardização. Aplicou-se quatro métodos de aglomeração Hierárquica: o *Complete Linkage*, o *Single Linkage*, o *Ward* e a Média Ponderada.

Grupos	%t	G4	G3	G2	G1
Grupo 4	57,14%	4	2	1	0
Grupo 3	96,30%	1	26	0	0
Grupo 2	100%	0	0	16	0
Grupo 1	50%	0	0	3	3
Total	87,50%	5	28	20	3

Tabela F.5.5.I.3 – Matriz de classificações

Utilizando os índices de atractividade associados aos segmentos masculino, feminino, dona de casa e não dona de casa para discriminar os quatro grupos chegou-se a uma matriz de classificação que nos dá 87,5% de observações correctamente classificadas (análise a posteriori). Os grupos com uma maior taxa de observações incorrectamente classificadas são os grupos de menor dimensão: o grupo 4 com 57% de acertos e o grupo 1 com 50% de acertos (Tabela F.5.5.I.3). As probabilidades *a priori* definiram-se, como sendo proporcionais ao tamanho dos grupos.

O p valor associado ao poder discriminante residual excluídas as duas primeiras funções é de 0,35.

A primeira função tem um coeficiente de correlação canónica de 0,81 e a segunda função de 0,61. A primeira função é responsável por 77% da variância explicada e a segunda por apenas 21%. Para medir a contribuição de uma variável para cada uma das funções utilizam-se os coeficientes estruturais. O modelo apresenta um Wilks lambda com o valor de 0,18. De acordo com os testes do χ^2 associados ao poder discriminante das funções, apenas são significativas as duas primeiras funções discriminantes. O p valor associado ao poder discriminante residual excluídas as duas primeiras funções é de 0,35. A primeira função tem um coeficiente de correlação canónica de 0,81 e a segunda de 0,61.

A primeira função é responsável por 77% da variância explicada e a segunda por apenas 21%. Para medir a contribuição de uma variável para cada uma das funções utilizam-se os coeficientes estruturais. Quando se tem múltiplas funções discriminantes interessa-nos uma medida global da importância da variável. O Índice de Potência representado pela expressão (1) é uma medida global da importância relativa de cada variável. Inclui a contribuição da variável para cada uma das funções e a contribuição da função para a solução global. Calcula-se como a soma do produto dos coeficientes estruturais ao quadrado pelo respectivo λ de cada função discriminante.

$$PI(X_j) = \sum_{i=1}^k [(\hat{\alpha}_{ji})^2 * \lambda_i / \sum \lambda_i] \quad (1)$$

$\hat{\alpha}_{ji}$ é o coeficiente estrutural da variável X_j para a função i Legenda: K é o número de funções discriminantes = min (g-1, p), g- n.º de grupos; p-nº de variáveis

Os Modelos de Exposição Necessários à Aquisição de Publicidade no Sector Televisivo

N.º	PROGRAMA
1	FILME DE SEGUNDA\DUelo IMORTAL
2	FILME DE SEGUNDA\GolPE QUASE PERFEITO
3	FILME DE SEGUNDA\OS INSPECTORES
4	FILME DE SEGUNDA\PAIXÃO SELVAJEM
5	FILME DE SEGUNDA\ROCKY V
6	FILME DE SEGUNDA\THELMA E LOUISE
7	FILME DE SEXTA\AMEÇA EXTRATERRESTRE
8	FILME DE SEXTA\DESPERTAR DO PASSADO
9	FILME DE SEXTA\ESQUECER PARIS
10	FILME DE SEXTA\STAR TREK
11	FILME DE TERÇA\CULPA FORMADA
12	FILME DE TERÇA\LIGAÇÃO YAKUZA
13	FILME DE TERÇA\OPERAÇÃO DE RISCO
14	FILME DE TERÇA\ROCKY IV
15	FILMES DE TERÇA\SILÊNCIO DE MORTE
16	FILMES DE TERÇA\VANDA
17	FILME DE QUARTA\SEM MEDO DE VIVER
18	FILME DE QUARTA\SEXO MENTIRA E OBSESSAO
19	FILME DE QUARTA\TEMPO DE DECIDIR
20	FILME DE QUARTA\VITIMAS DA JUSTIÇA
21	FILME DE QUINTA\ANGIE - UMA MULHER SÓ
22	FILME DE QUINTA\RISCO CALCULADO
23	FILME ESP DE SAB\HARRY POTTER E A PEDRA FILOSOFAL
24	MATINE DEMINGO I\007 - ALVO EN MOVIMENTO
25	MATINE DOMINGO I\BRAVEHEART – O DESAFIO DO GUERREIRO
26	MATINE DOMINGO I\CIDADE EM PERIGRO
27	MATINE DOMINGO I\HELENA DE TROIA
28	MATINE DOMINGO I\MISSÃO IMPOSSIVEL
29	MATINE DOMINGO II\GODZILLA
30	MATINE DOMINGO II\JUSTIÇA VERMELHA
31	MATINE DOMINGO II\TUBARÕES
32	MATINE DOMINGO II\PERFUME DE MULHER
33	MATINE DOMINGO III\LIMITES VERTICAIS
34	MATINE SABADO I\007 – MISSÃO ULTRA-SECRETA
35	MATINE SABADO I\COMANDO
36	MATINE SABADO I\COMO AGARRAR UM MARIDO
37	MATINE SABADO I\O REGRESSO DO PESTINHA
38	MATINE SABADO I\OPERAÇÃO FLECHA QUEBRADA
39	MATINE SABADO II\007 - GOLDENEYE
40	MATINE SABADO II\007 – O AMAÑHA NUNCA MORRE
41	MATINE SABADO II\LA MUMIA
42	MATINE SABADO II\O REGRESSO DA MUMIA
43	MATINE SABADO II\ROMANCE PERIGOSO
44	FILME DO DOMINGO \ALVO DE TRAIÇÃO
45	FILME DO DOMINGO \BERMUDAS - TRIÂNGULO MALDITO
46	FILME DE DOMINGO \CÓDIGO VERMELHO
47	FILME DE DOMINGO \MEDO DE MORTE
48	FILME DE DOMINGO \RISCO CALCULADO
49	FILME DE DOMINGO \UM VIAJEM SEM FIM
50	FILME DE SABADO \CAÇA AO HOMEM
51	FILME DE SABADO \FENÓMENO
52	FILME DE SÁBADO\PLAYERS CLUB
53	FINAL TAÇA DE PORTUGAL
54	FUTBOL - TAÇA DE PORTUGAL\BENFICA X FC PORTO
55	LISBOA REGRESSA AO PARQUE (R)
56	PROCISSÃO DAS VELAS
57	VIVA O ZOO

Quadro F.5.5.I.2 – Associação de programas a faixas etárias de acordo com a aplicação de metodologias de MDS. Legenda - Idades 35/44, Idades 45/54, Idades 55/64, Idades +64, Idades 4/14, Idades 15/24, Idades 25/34

ID	PROGRAMAS	TIPOLOGIAS
1	FILME DE SEGUNDA \DUELO IMORTAL	Acção
2	FILME DE SEGUNDA \GOLPE QUASE PERFEITO	Suspense
3	FILME DE SEGUNDA \OS INSPECTORES	Suspense
4	FILME DE SEGUNDA \PAIXÃO SELVAGEM	Suspense
5	FILME DE SEGUNDA \ROCKY V	Acção
	FILME DE SEGUNDA \THELMA E LOUISE	Drama
6		
7	FILME DE SEXTA \AMEAÇA EXTRATERRESTRE	Ficção Científica
8	FILME DE SEXTA \DESPERTAR DO PASSADO	Drama
9	FILME DE SEXTA \ESQUECER PARIS	Comédia
10	FILME DE SEXTA \STAR TREK - GERAÇÕES	
		Ficção Científica
11	FILME DE TERÇA \CULPA FORMADA	Suspense
12	FILME DE TERÇA \LIGAÇÃO YAKUZA	Acção
13	FILME DE TERÇA \OPERAÇÃO DE RISCO	Acção
14	FILME DE TERÇA \ROCKY IV	Acção
15	FILME DE TERÇA \SILÊNCIO DE MORTE	Suspense
16	FILME DE TERÇA \UM PEIXE CHAMADO VANDA	
		Comédia
17	FILME DE QUARTA \SEM MEDO DE VIVER	Drama
18	FILME DE QUARTA \SEXO MENTIRA E OBSSessão	Drama
19	FILME DE QUARTA \TEMPO DE DECIDIR	Drama
20	FILME DE QUARTA \VITIMAS DA JUSTIÇA	Drama
21	FILME DE QUINTA \ANGIE - UMA MULHER SÓ	Drama
22	FILME DE QUINTA \RISCO CALCULADO	Suspense
23	FILME DE SABADO \HARRY POTTER E A PEDRA FIOSOFAL	Fantasia
24	MATINE DOMINGO I \007 - ALVO EM MOVIMENTO	Acção
25	MATINE DOMINGO I \BRAVEHEART - O DESAFIO DO GUERREIRO	Acção
26	MATINE DOMINGO I \CIDADE EM PERIGRO	Aventura
27	MATINE DOMINGO I \HELENA DE TROIA	Histórico/Épico
28	MATINE DOMINGO I \MISSÃO IMPOSSIVEL	Suspense
29	MATINE DOMINGO II \GODZILLA	Suspense
30	MATINE DOMINGO II \JUSTIÇA VERMELHA	Drama
31	MATINE DOMINGO II \TUBARÕES	Suspense
32	MATINE DOMINGO II \PERFUME DE MULHER	Drama
33	MATINE DOMINGO III \LIMITE VERTICAL	Suspense
34	MATINE SABADO I \007 – MISSÃO ULTRA-SECRETA	Acção
35	MATINE SABADO I \COMANDO	Acção
36	MATINE SABADO I \COMO AGARRAR UM MARIDO	Comédia
37	MATINE SABADO I \O RETORNO DO PESTINHA	Comédia
38	MATINE SABADO I \OPERAÇÃO FLECHA QUEBRADA	Suspense
39	MATINE SABADO II \007 - GOLDENEYE	Acção
40	MATINE SABADO II \007 – O AMAÑHA NUNCA MORRE	Acção
41	MATINE SABADO II \A MUMIA	Aventura
42	MATINE SABADO II \O RETORNO DA MUMIA	Aventura
43	MATINE SABADO II \ROMANCE PERIGOSO	Suspense
44	FILME DE DOMINGO \ALVO DE TRAIÇÃO	Aventura
45	FILME DE DOMINGO \BERMUDAS - TRIÂNGULO MALDITO	Suspense
46	FILME DE DOMINGO \CÓDIGO VERMELHO	Acção
47	FILME DE DOMINGO \MEDO DE MORTE	Suspense
48	FILME DE DOMINGO \RISCO CALCULADO	Suspense
49	FILME DE DOMINGO \UMA VIAGEM SEM FIM	
		Terror
50	FILME DE SABADO \CAÇA AO HOMEN	Acção
51	FILME DE SABADO \FENÓMENO	Drama
52	FILME DE SABADO \PLAYERS CLUB	Drama
53	FINAL TAÇA DE PORTUGAL	NULL
54	FUTEBOL - TAÇA DE PORTUGAL \BENFICA X FC PORTO	Futebol
55	LISBOA REGRESSA AO PARQUE (R)	Revista
56	PROCISSÃO DAS VELAS	NULL
57	VIVA O ZOO	NULL

Quadro F.5.5.I.3 – Identificação de programas utilizados na análise de Cluster

Variáveis	Função 1	Função 2	Função 3
Masculino	-0,78	-0,49	0,12
Feminino	0,61	0,43	-0,28
Dona de Casa	0,69	-0,63	-0,36
Não Dona de Casa	-0,71	0,70	0,10

Tabela F.5.5.I.4 – Coeficientes estruturais das variáveis com as funções discriminantes

Neste caso a variável masculino apresenta um índice de potência no valor de 0,52, o feminino de 0,33, dona de casa de 0,45 e não dona de casa de 0,49. Os índices apresentam valores semelhantes, excluindo o sexo feminino, por conseguinte, esta será a variável que se diferencia em menor grau entre os vários conjuntos de programas. O sexo masculino e todos os telespectadores que não são responsáveis pelas compras do agregado familiar são os segmentos que melhor diferenciam os grupos. Com base no gráfico que nos permite visualizar as pontuações que foram calculadas com base nas funções que se definem pelos coeficientes não estandardizados verificamos que para as duas primeiras funções discriminantes o grupo 4 corresponde ao que apresenta uma maior dispersão.

Ao grupo 2 associamos maiores índices de atractividade ao segmento masculino e a quem não representa as donas de casa e ao grupo 3 associamos maiores índices ao segmento feminino e às donas de casa. Ao grupo 4 associamos maiores índices de atractividade às donas de casa e ao sexo masculino e ao grupo 1 todos os telespectadores que não são donas de casa e o sexo feminino.

Portanto, por último considerando os segmentos feminino e masculino, os homens apresentam maiores índices de atractividade para os filmes do Domingo à tarde da tipologia suspense, por filmes do Sábado à tarde de acção e pelo futebol. Para o sexo feminino em geral destacam-se os filmes em dias laborais da tipologia drama, Ficção Científica e ao Sábado de acção (talvez devido ao denominado “co-viewing”). Ao segmento masculino e aos telespectadores que não são donas de casa associamos os filmes de fim-de-semana sobretudo os espaços da tarde e das tipologias de acção, suspense e o futebol. Ao sexo feminino e às donas de casa associamos os filmes ao longo de toda a semana e de todas as tipologias. Com base nos resultados as associações parecem ser sobretudo função da disponibilidade do público e não de preferências por tipologias cinematográficas.

Ib. Resultados para o canal SIC e para programas que apresentam vários episódios

Para a análise dos índices de atractividade do canal SIC referentes a programas com mais de um episódio recorreremos à análise em Componentes Principais Simbólica que procura a melhor representação num espaço factorial de uma base de dados simbólica descrita por variáveis da tipologia intervalo. Recorreremos a três abordagens: SPCA Vértices; RT-SPCA e a uma Estratégia Mista. No método dos vértices, o critério a maximizar consiste na variância entre objectos simbólicos em vez da variância total. A abordagem RT-PCA traduz-se numa transformação da ordenação das descrições simbólicas por forma a colocar em evidência as diferenças existentes entre os vários objectos simbólicos.

Geometricamente consiste em realizar a translação dos vértices menores dos objectos simbólicos para a origem do espaço factorial e realizar a PCA sobre os vértices maiores. A comparação entre objectos simbólicos é realizada atendendo à forma e ao tamanho das suas representações no espaço factorial. Neste caso pretende-se uma transformação da amplitude dos intervalos, por forma a analisar a forma e o tamanho dos objectos independentemente da sua posição.

A estratégia mista consiste em poder representar os SO no espaço factorial, projectando os vértices dos hipercubos nos eixos factoriais obtidos segundo o algoritmo RT_PCA. Assim, tenta tirar partido de ambas as metodologias por forma a melhorar a qualidade de representação atendendo as diferenças estruturais e de forma dos SO.

As variáveis neste caso correspondem aos vários índices da atractividade e os SO aos programas com mais de um episódio, obtendo-se, assim, um intervalo de atractividade para cada um dos segmentos. Pretendemos associar programas a determinados segmentos sociodemográficos em função da atractividade exercida. Utilizam-se os três algoritmos e comparam-se os resultados alcançados.

Os SO são representados no plano factorial por rectângulos com lados paralelos aos eixos. Como saída aos algoritmos obtemos as coordenadas dos vértices dos hipercubos originais no plano factorial (“SO convex hull plot”) e as coordenadas dos extremos dos intervalos dos rectângulos que reproduzem a projecção dos hipercubos nos eixos factoriais (“SO rectangular plot”), ou seja, a menor área rectangular que contem a projecção do hipercubo original no espaço factorial.

Como resultado gráfico obtemos, também, os círculos de correlação das variáveis com os eixos factoriais e os intervalos de correlação das variáveis com os eixos.

Assim como se verifica na PCA clássica o algoritmo dá-nos os valores absolutos e a percentagem de inércia explicada pelos eixos; alguns índices para avaliar a qualidade da representação dos SO no plano factorial e a contribuição de cada objecto e de cada variável para a obtenção dos eixos.

➤ Algoritmo Misto

Complementando as análises com a realização de uma análise ACP simbólica (algoritmo Misto) aos programas da SIC e considerando a totalidade dos segmentos é necessário reter cinco eixos factoriais para conseguir uma inércia explicada acumulada de 79%. Ao primeiro eixo corresponde 33%, ao segundo 16,87%, ao terceiro 14,2%, ao quarto 9,2% e ao quinto o valor de 6,03%.

Segundo a análise ACP os espaços informativos da noite podem-se associar às idades compreendidas entre os 25 e os 34 anos e entre os 45 e os 54 anos e aos lares que não têm TV cabo. Ao espaço informativo da tarde de fim-de-semana associam-se as faixas etárias dos 4/14 e dos 15/24, a região de Lisboa, a classe D e os homens. Ao espaço informativo da tarde de Segunda a Sexta associam-se os segmentos: 15/34, a Classe A/B, Lisboa, Costa Norte e as não donas de casa. Segundo a análise ACP as telenovelas de pré-prime-time associam-se às idades compreendidas entre os 35 e os 54 anos, a todas as classes sociais com excepção da mais baixa e aos lares que não têm TV cabo.

As restantes telenovelas da noite associam-se a todas as classes, aos telespectadores de idades compreendidas entre os 25 e os 34 anos e 45 a 54 anos e aos lares sem TV cabo. O espaço informativo da noite de fim-de-semana associa-se a todas as regiões e classes com excepção da classe D, faixas etárias 25-34 e 45-54 e lares com TV cabo. O espaço informativo da tarde não se associa a nenhum factor, nem o espaço informativo da noite em dias semanais.

O *Gráfico F.5.5.1.5* representa os intervalos de correlações das variáveis com os eixos factoriais segundo o algoritmo dos vértices e segundo o algoritmo da estratégia mista. Segundo ambos os algoritmos verificamos que o 1ºQ está associado à Classe A/B e ao sexo feminino; o 2º Q a todas as faixas etárias, à situação no lar e aos lares c/s TV cabo; ao 4º Q à classe C1, C2, D, masculino, Costas, Interior e o Sul.

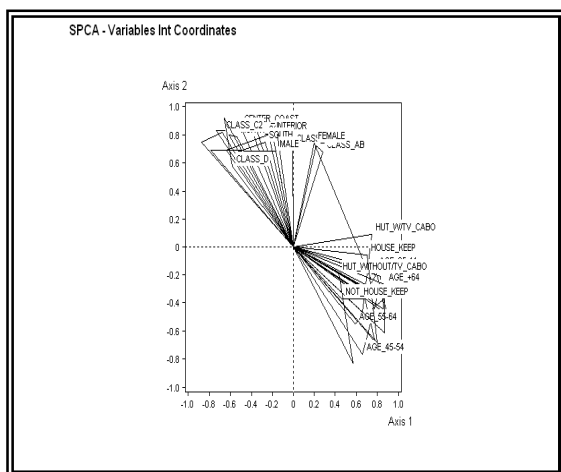


Gráfico F.5.5.I.5 – Intervalo de correlações das variáveis segundo o algoritmo dos vértices

Programas que apresentam múltiplos episódios – Análises MDS

➤ **Classes Sociais**

Segundo as análises de Regressão Linear a primeira dimensão MDS está associada à Classe A/B ($\beta=-0.88$) e à Classe D ($\beta=0.71$). A segunda dimensão associa-se à Classe C1 ($\beta=-0.90$), à Classe C2 ($\beta=0.54$) e à Classe D ($\beta=0.159$). A A/B($\beta=0.27$), à Classe C2 ($\beta=-0.68$) e à Classe D ($\beta=0.49$) (*Gráfico F.5.5.I.6*).

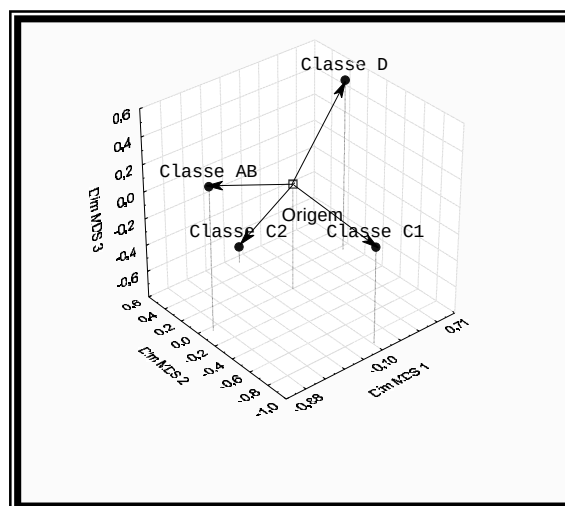


Gráfico F.5.5.I.6 – Classes sociais no espaço MDS

À classe C1 associam-se os filmes ao longo de toda a semana e de todas as tipologias, mas com uma maior incidência nas tipologias de drama, acção e comédia. À classe A/B podemos associar os filmes do final da noite em dias laborais a partir das 23H. A selecção parece ser sobretudo função da disponibilidade do público e em menor grau da tipologia, uma vez que se encontram distribuídos por todas as tipologias. À classe D associa-se-lhe filmes de fim de semana, ao longo de toda a tarde ao Sábado e ao Domingo e da Sexta à noite.

➤ As faixas etárias

Realizando uma análise idêntica para as faixas etárias, segundo a RLM a primeira dimensão MDS está associada às faixas etárias 4/14 ($\beta=-0.71$), 25/34 ($\beta=0.42$) e +64 ($\beta=-0.22$). A segunda dimensão está associada aos segmentos 25/34 ($\beta=-0.34$) e 35/44 ($\beta=0.85$). A terceira dimensão aos segmentos 4/14 ($\beta=0.23$), 15/24 ($\beta=-0.34$), 25/34 ($\beta=0.63$), 35/44 ($\beta=0.19$) e +64 ($\beta=-0.67$).

Ao público de idades compreendidas entre os 4 e os 14 anos associam-se os filmes ao longo de toda a semana e aos fins de semana também à tarde e sobretudo da tipologia comédia. Para a faixa etária dos 15 aos 24 anos e dos 25 aos 34 anos associam-se os filmes ao longo de toda a semana e dos fins de semana à tarde e de todas as tipologias. A faixa etária dos 35/44 anos apresenta maiores índices de atractividade por filmes ao longo de toda a semana e aos fins de semana no horário da noite e sobretudo da tipologia acção. Aos telespectadores de idades superiores aos 64 anos associam-se sobretudo os filmes da noite.

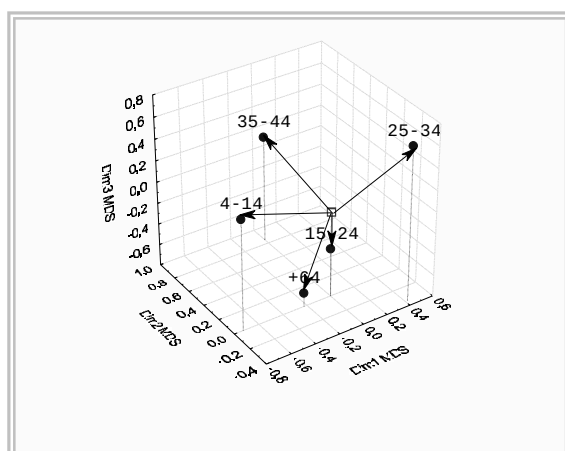


Gráfico F.5.5.I.7 – Variáveis Sig. Na Regressão (faixas etárias)

➤ Sexo

Ao sexo masculino associa-se, sem dúvida, os filmes de acção da noite sobretudo aos Sábados e em segundo lugar as comédias. Ao sexo feminino associam-se os filmes em especial das tipologias família e de comédia aos fins de semana à tarde. As donas de casa apresentam maiores índices de atractividade por filmes das tipologias de drama e de suspense, em todos os horários e repartidos por todos os dias.

➤ Dona de Casa

Ao público que não é responsável pelas compras do lar associam-se os filmes ao longo de toda a semana, especialmente os que são transmitidos no período da noite e de todas as tipologias programáticas (Tabela F.5.5.I.5).

Variáveis	Função 1	Função 2	Função 3
Masculino	0,81	-0,86	0,26
Feminino	0,82	0,88	-0,23
D Casa	0,64	0,74	0,30
ND Casa	0,63	-0,73	-0,26

Tabela F.5.5.I.5 – Coeficientes das variáveis sexo e situação no lar

Ic. Resultados para o canal RTP1

O canal estatal é o que apresenta um menor número de espaços de programação destinados à transmissão de filmes. Para associar filmes a segmentos com base nos seus índices de atractividade realizou-se uma análise de *clusters* segundo os algoritmos de agrupamento hierárquicos.

□ **Programas com um único episódio**

Não foi possível segundo estas metodologias definir agrupamentos bem definidos que permitam uma explícita diferenciação das variáveis em consideração. Alternativamente recorreu-se ao algoritmo K Média. Experimentando várias configurações iniciais, o melhor resultado está associado a cinco agrupamentos. O primeiro agrupamento está constituído por filmes transmitidos durante a semana (15 observações) e de fim-de-semana (9 observações) essencialmente da noite e na maioria associados à tipologia drama (10) e de suspense (4).

O segundo agrupamento inclui filmes dos dias laborais (10 observações) e de fim-de-semana (8 observações) repartidos pela tarde e noite e na sua maioria da tipologia comédia (6) e de acção (4). O terceiro *cluster* corresponde a filmes da semana e da tipologia drama. O quarto agrupamento associa-se a filmes da semana e das tipologias drama, suspense, comédia, acção e guerra e o quinto agrupamento a filmes da semana e das tipologias comédia, acção, drama e suspense.

□ **Programas com mais de um episódio**

Por observação das Zoom Stars para espaços agregados de transmissão de filmes e dos agrupamentos que resultam da aplicação do K Média concluímos que atendendo à classe social A/B são maiores os índices de atractividade associados aos filmes da noite ao longo de toda a semana e de todas as tipologias. Os seus índices são muito elevados, portanto a classe AB parece seleccionar em maior grau a RTP1 em detrimento dos canais privados. Para as classes C1 e C2 os índices são menores, no entanto também estão

associadas aos filmes da noite ao longo de toda a semana. Os índices da classe D são muito reduzidos para este canal. A menor classe social selecciona essencialmente os filmes dos canais privados e a classe mais alta os programas da RTP. As restantes duas classes sociais parecem seleccionar filmes do canal RTP1 de Segunda a Sexta, independentemente da tipologia e aos fins de semana existe uma maior exposição aos programas dos canais privados.

Atendendo às faixas etárias o público com idades inferiores aos 14 anos apresenta uma exposição quase que inexistente aos espaços de transmissão de filmes deste canal. A audiência com idades compreendidas entre os 15 e os 24 anos seleccionam ligeiramente mais estes programas, especialmente à noite em dias laborais e sobretudo das tipologias drama e comédia. As faixas etárias dos 25 aos 34 anos apresentam maiores índices em relação à faixa etária anterior, mas, também, apenas para os filmes da noite em dias laborais e de todas as tipologias.

As faixas etárias entre os 35 e os 54 anos seleccionam frequentemente os filmes deste canal do período da noite. Para o segmento 55/64 anos constata-se uma grande redução nos índices de atractividade e, por último para os telespectadores que superam os 64 anos verifica-se uma exposição ligeiramente superior, não só no período da noite, mas também no decorrer das tardes ao fim-de-semana.

Portanto, o factor idade parece contribuir para um acréscimo na selecção de filmes deste canal (com excepção do segmento 55/64 anos onde se verifica uma inversão nesta tendência), mas sobretudo no decorrer da noite e em dias laborais.

O sexo masculino apresenta índices elevados, independentemente do dia, hora e tipologia. Aparentemente seleccionam em maior grau a RTP1 do que os canais privados e para as mulheres verifica-se o inverso. Verifica-se a mesma tendência com as donas de casa e os telespectadores que não são donas de casa. Os telespectadores que são responsáveis pelas compras do lar não seleccionam este canal de forma significativa e os telespectadores que o não são coincidem com o padrão associado e descrito para o sexo masculino. O público com TV Cabo apresenta reduzidos índices e os lares que não têm TV cabo, quando seleccionam este canal é sobretudo no período da noite.

Portanto, no que diz respeito aos filmes verificamos que o consumo é bastante aleatório e, logo, não é possível realizar uma nítida associação dos suportes aos vários segmentos em função das tipologias, dias e horários de transmissão. A exposição é superior ao longo da semana, ou seja, no período em que não existe uma contra-programação, mas, sim, uma estruturação das grelhas concorrentes por complementaridade. As excepções são o sexo masculino, os mais idosos e a classe social mais alta.

Para complementar as análises procurou-se a associação de programas aos factores segundo o algoritmo do Mix-Pca (ACP simbólica) e segundo o método dos Vértices. Para ambos os algoritmos é necessário reter cinco eixos para alcançarmos uma variância explicada superior aos 75%. Estes resultados resultam da análise dos coeficientes de correlação das variáveis com os eixos factoriais e dos índices que traduzem a qualidade de representação dos SO pelos diferentes eixos. A representação gráfica dos “SO” no plano definido pelos dois principais eixos factoriais não produz uma

clara diferenciação dos diversos programas, existindo grandes sobreposições. Por exemplo, ao espaço informativo da tarde de fim-de-semana e aos principais concursos do canal estatal associa-se segundo ambos os algoritmos a Classe A/B, C1, o sexo masculino, a faixa etária dos 15-24 e os telespectadores que não são donas de casa. Segundo o algoritmo dos vértices o espaço informativo da noite associa-se a Lisboa, Costa Norte e ao Porto, às classes C1 e C2 e à faixa etária dos 15 aos 24 anos. Face aos resultados pouco claros decorrentes da aplicação da análise ACP, realizou-se também uma análise MDS simbólica (Interscal).

A análise MDS simbólica permite representar os intervalos de dissimilaridades entre pares de objectos num espaço de menor dimensionalidade. O objectivo das análises corresponde a detectar relações entre os objectos e descobrir quais são as dimensões que dão origem ao espaço multidimensional. Uma vez mais existe uma grande concentração de objectos simbólicos perto da origem do espaço MDS, não produzindo bons resultados. Concluindo, verificamos que para a maioria dos programas do canal estatal RTP1, sejam eles filmes ou não, existe uma maior exposição ao sexo masculino, aos telespectadores que não são responsáveis pelas compras no lar, às classes sociais mais elevadas e sobretudo às regiões do Norte de Portugal Continental.

Ao nível das faixas etárias, verificamos que a exposição é mais reduzida na globalidade ao público mais jovem. Seleccionam os programas juvenis e em relação aos filmes os mais jovens até aos 14 anos apenas algum filme do fim-de-semana e se considerarmos até aos 24 anos apenas algum filme de Segunda a Sexta.

As faixas etárias até aos 24 anos apresentam os menores valores de índices de atractividade. O sexo masculino e a classe AB seleccionam todo o tipo de programas. Podemos associar à classe social D algum filme de comédia e de acção ao longo de toda a semana e o concurso “O Preço Certo”. Em relação aos filmes, a partir dos 64 anos existe uma exposição ao longo de toda a semana e a todas as tipologias e a partir dos 55 anos sobretudo das tipologias de comédia e de acção.

II. Análises aos índices de Adesão

Segmentos	Programas com Menor Índice de Adesão
Classe Social A/B	Concursos, telenovelas, talk-shows e espaços de informação à tarde e programas matinais
Classe Social C1	Concursos, telenovelas, talk-shows e espaços de informação à tarde e programas matinais
Classe Social C2	Concursos, telenovelas, talk-shows e espaços de informação à tarde e programas matinais
Classe Social D	Filmes à noite toda a semana e à tarde também aos fins-de-semana
Masculino	Programas juvenis
Feminino	Filmes à noite toda a semana e à tarde também aos fins-de-semana
Idades 4-14	Concursos, telenovelas, talk-shows e espaços informativos à tarde, programas matinais e filmes à noite de Segunda a Sexta
Idades 15-24	Programas juvenis
Idades 25-34	Concursos, telenovelas, talk shows e espaços informativos à tarde e programas matinais
Idades 35-44	Concursos, telenovelas, talk shows e informativos à tarde e programas matinais
Idades 45-54 / 55-64	Programas juvenis
Idades + 64	Filmes à noite toda a semana e à tarde também os fins-de-semana

Quadro F.5.5.II.1 – Programas do canal RTP1 com maiores e menores índices de adesão

Segmentos	Programas com maior Índice de Adesão
Classe Social A/B	Filmes à noite das Quartas, Quintas e Sextas, séries, formula1
Classe Social C1	Filmes à noite das quartas, quintas e sextas, séries, formula1
Classe Social C2	Filmes toda a semana e do fim-de-semana à noite e à tarde, séries
Classe Social D	Concursos, telenovelas, talk shows e programas de informação à tarde e programas matinais
Masculino	Filmes à noite toda a semana e à tarde também aos fins-de-semana
Feminino	Concursos, telenovelas, talk -hows e programas de informação à tarde e programas matinais
Idades 4-14	Programas juvenis
Idades 15-24	Filmes à noite das quartas, quintas e sextas, séries, formula1
Idades 25-34	Filmes à noite das quartas, quintas e sextas, séries, formula1
Idades 35-44	Filmes à noite toda a semana e à tarde também aos fins-de-semana
Idades 45-54	Filmes à noite das quartas, quintas e sextas, séries, formula1, espaços informativos da noite, futebol, debate, galas, concursos, comédia
Idades 55-64	Espaços informativos da noite, futebol, debate, galas, concursos e comédia
Idades + 64	Concursos, telenovelas, talk-shows e programas de informação à tarde e programas matinais

Quadro F.5.5.II.2 – Programas do canal RTP1 com maiores e menores índices de adesão

Considerando os índices de adesão associados aos segmentos sóciodemográficos tentou-se identificar as tipologias programáticas que maximizam estes índices, recorrendo a metodologias de Regras de Associação e de *Clustering* simbólicas. Para alguns dos segmentos aplicaram-se as Regras de Associação obtendo-se gráficos que ilustram as Regras e as ligações existentes entre os diversos programas atendendo à frequência com que ocorrem simultaneamente nas bases de dados com valores de índices de afinidade superiores à unidade. Com base nas Regras derivadas realizaram-se as seguintes associações, segundo constam do seguintes Quadros F.5.5.II.1-4.

SEGMENTOS	PROGRAMAS COM MAIORES ÍNDICES DE ADESÃO	PROGRAMAS COM MENORES ÍNDICES DE ADESÃO
Classe Social A/B	Filmes ao longo de toda a semana à tarde e à noite	Telenovelas e séries à tarde, muitas transmitidas por segunda vez, espaço informativo da tarde, programas juvenis
Classe Social C1	Filmes ao longo de toda a semana à tarde e à noite	Telenovelas e séries à tarde, muitas transmitidas por segunda vez, espaço informativo da tarde, programas juvenis
Classe Social C2	Filmes ao longo de toda a semana à tarde e à noite	Telenovelas e séries à tarde, muitas transmitidas por segunda vez, espaço informativo da tarde, programas juvenis
Classe Social D	Telenovelas e séries à tarde, muitos transmitidos por segunda vez, espaço informativo da tarde, programas juvenis	Filmes toda a semana à tarde e à noite
Masculino	Filmes toda a semana à tarde e à noite	Telenovelas e séries à tarde, muitas transmitidas por segunda vez, espaço informativo da tarde, programas juvenis
Feminino	Telenovelas e séries à tarde, muitas transmitidas por segunda vez, espaço informativo da tarde	Filmes toda a semana à tarde e à noite
Idades 4-14	Programas juvenis, séries, telenovelas de prime time, espaços informativos da noite, filmes fim-de-semana à tarde	Telenovelas e séries à tarde, muitas transmitidas por segunda vez, espaço informativo da tarde, programas juvenis
Idades 15-24	Filmes toda a semana à tarde e à noite	Programas juvenis, séries, telenovelas de prime time, espaço informativo da noite, filmes fim-de-semana à tarde
Idades 25-34	Filmes toda a semana à tarde e à noite	Programas juvenis, séries, telenovelas de prime time, espaços informativos da noite, filmes fim-de-semana à tarde
Idades 35-44	Filmes toda a semana à tarde e à noite	Telenovelas e séries à tarde, muitas transmitidas por segunda vez, espaço informativo da tarde, programas juvenis
Idades 45-54	Filmes toda a semana à tarde e à noite, telenovelas de prime time, séries, espaço informativo da noite	Telenovelas e séries à tarde, muitas transmitidas por segunda vez, espaço informativo da tarde, programas juvenis
Idades 55-64	Telenovelas e séries todo o dia, espaços informativos da tarde e noite	Filmes toda a semana à tarde e à noite
Idades + 64	Telenovelas e séries à tarde, muitas transmitidas por segunda vez, Jornal da Tarde	Filmes toda a semana à tarde e à noite

Quadro F.5.5.II.3 – Programas do canal TVI com maiores e menores índices de adesão

SEGMENTOS	PROGRAMAS COM MAIOR ÍNDICE DE ADEÇÃO	PROGRAMAS COM MENOR ÍNDICE DE ADEÇÃO
Classe Social A/B	Filmes e séries à noite e programas culturais	Programas matinais e da tarde (talk-shows), telenovelas a longo de todo o dia, comédias e filmes da tarde aos fins-de-semana
Classe Social C1	Filmes e séries à noite, programas culturais, galas	Telenovelas de prime time, filmes à tarde, espaços informativos, comédias
Classe Social C2	Programas Juvenis	Telenovelas de prime time, filmes à tarde, espaços informativos comédias
Classe Social D	Espaço informativo da Noite, telenovelas de prime time, programas matinais e das tardes (talk-show)	Filmes e Séries à noite e programas culturais
Masculino	Filmes e Séries à noite e programas culturais	Jornal Noite, telenovelas de prime-time, programas matinais e das tardes (talk-show)
Feminino	Jornal Noite, telenovelas de prime time, programas matinais e das tardes (talk-show)	Filmes e Séries à noite e programas culturais
Idades 1-14	Programas juvenis	Filmes e Séries à noite e programas culturais
Idades 15-24	Filmes e Séries à noite	Juvenil
Idades 25-34	Filmes e Séries à noite	Juvenil, Jornal Noite, telenovelas de prime time, programas matinais e das tardes (talk-show)
Idades 35-44	Filmes e Séries à noite	Juvenil, Jornal Noite, telenovelas de prime time, programas matinais e das tardes (talk-show)
Idades 45-54	Filmes e Séries à noite	Juvenil
Idades 55-64	Jornal noite e tarde, telenovelas de prime-time e das tardes, programas matinais e das tardes (talk-show) e filmes da tarde	Juvenil
Idades + 64	Jornal Noite, telenovelas de prime time, programas matinais e das tardes (talk-show)	Filmes e Séries à noite e programas culturais e juvenis

Quadro F.5.5.II.4 – Programas do canal SIC com maiores e menores índices de Adesão

III. Análises aos Índices de Afinidade

Os índices de afinidade resultam do quociente entre os valores de *ratings* associados a segmentos sóciodemográficos e os *ratings* do segmento de referência (considerou-se o universo). Utilizaram-se metodologias de *Data Mining* e Simbólicas, nomeadamente a análise ACP e os algoritmos de *clustering*.

A base de dados simbólica corresponde à que agrega em variáveis da tipologia intervalo os índices de afinidade dos vários episódios de um programa para os três principais canais generalistas. Diferenciam-se os resultados em função dos seguintes segmentos: regiões, classes, faixas etárias, sexo e ser responsável pelas compras no lar.

Iniciando as análises com o segmento regiões geográficas aplicou-se a análise ACP simbólica para avaliar o grau de correlação existente entre os índices associados às diferentes regiões. Pretende-se, também, identificar os programas que apresentam maiores valores de índices de afinidade para cada um dos segmentos considerados.

Verificamos que para três eixos factoriais a percentagem de inércia explicada é de 80% (*Tabela F.5.5.III.1*). Os segmentos mais correlacionadas com o primeiro factor são as regiões de Lisboa (0,94) e o LC (0,86) e em menor grau o Interior (0,55). As variáveis que apresentam uma maior correlação com o segundo factor são o LN (0,95) e o Porto (0,80) e o mais penalizado é o Sul (0,65). Por último com o terceiro factor o LC (0,87) apresenta maiores valores e os restantes segmentos uma ordem de grandeza muito semelhante. Conforme se observa nos seguintes gráficos (*Gráficos F.5.5.III.1/2*) as várias regiões apresentam valores muito próximos, ou seja, a evolução destes índices não é significativamente diferenciada atendendo a esta variável de segmentação do painel. O primeiro factor associado a Lisboa e Litoral Centro, apresenta maiores contribuições absolutas para programas do canal SIC, por exemplo, para os espaços de filmes; atendendo aos espaços informativos apenas para o Primeiro Jornal (S) e para as séries/telenovelas salientam-se os seguintes programas: O programa matinal em dias laborais, a segunda telenovela da noite e o programa de humor do horário nobre ao fim-de semana.

	Valor P.	% Exp.	% Exp. Ac		
Ev.1	0.52835	37.40853	37.40853		*****
Ev.2	0.36946	26.15860	63.56713		*****
Ev.3	0.23289	16.48935	80.05648		*****
Ev.4	0.13460	9.53005	89.58653		***
Ev.5	0.12060	8.53847	98.12500		***
Correlações					
Var.	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Lis	0.94461	0.74755	0.80001	0.82160	0.78177
Porto	0.80928	0.81683	0.77181	0.85908	0.83182
LN	0.75468	0.95431	0.81425	0.81961	0.87300
LC	0.86540	0.72266	0.87578	0.81401	0.83702
Int	0.55188	0.74298	0.83534	0.81548	0.77784
Sul	0.71579	0.65750	0.77328	0.82590	0.87142

Tabela F.5.5.III.1 – Descrição de factores

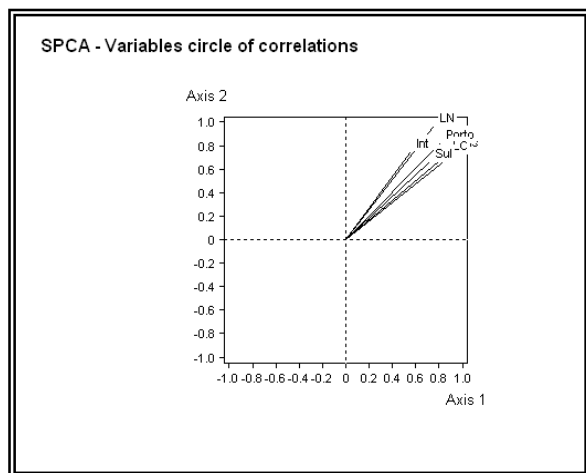


Gráfico F.5.5.III.1 – Correlação entre índices de afinidade para as regiões geográficas

O terceiro factor apresenta valores de correlação semelhantes para as várias regiões e maiores contribuições por parte de programas que não apresentam uma maior incidência em determinado canal e onde constam todas as tipologias. Com base nestas análises não foi possível fazer associações de grande relevância.

Se considerarmos as classes sociais (Gráfico F.5.5.III.2) e, também, por aplicação da ACP simbólica com dois eixos factoriais obtemos 84% de inércia explicada. A classe D apresenta maior correlação com o primeiro eixo (0,95); a classe C2 com o segundo eixo (0,94) e o terceiro com a classe AB (0,87).

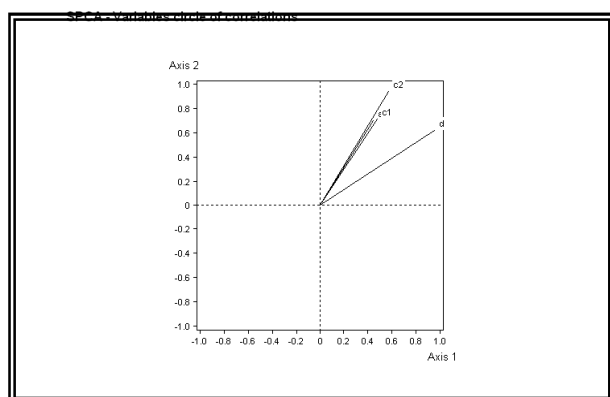


Gráfico F.5.5.III.2 – Correlação entre índices de afinidade para as classes sociais

Para a TVI, também, incide em filmes (matinés de Domingo, Filmes Quinta, Sábado e Sexta), em algumas séries e no espaço informativo da tarde (FS). Para a RTP1 sobressai o espaço informativo da tarde ao longo de toda a semana, filmes da noite e o programa matinal. Portanto, corresponde essencialmente aos espaços informativos da tarde, filmes em todos os canais, dias e horários, a ficção nacional, o talk-show matinal, entre outros. O segundo factor correlacionado com o LN está em maior grau associado aos programas da TVI e das tipologias de séries, telenovelas e talk-shows.

Com base no círculo de correlações verificamos que as classes C1 e C2 estão muito correlacionadas e que a classe D corresponde à que mais se distânciava das restantes. O terceiro factor associado à classe AB associa-se essencialmente a filmes de todos os canais, dias e períodos horários e a alguns programas. Apresenta contribuições mais reduzidas e menor número de programas com valores mais elevados. Os filmes estão distribuídos ao longo de toda a semana, canais e períodos horários, existindo ao fim-de-semana maiores índices para o Domingo à tarde. O primeiro factor associado à classe D apresenta um elevado número de programas de todos os canais e tipologias e com razoáveis valores de contribuições. Os programas mais correlacionados da RTP são os seguintes: o “Preço Certo”, o concurso da noite, filmes da RTP da noite ao longo de toda a semana, talk-show da tarde e da manhã, espaço informativo da tarde ao longo de toda a semana. Corresponde a filmes, concursos, programas adjacentes aos espaços informativos, jornal da tarde, talk.show da tarde e ao programa matinal. Para o canal SIC os maiores valores associam-se aos programas que correspondem a filmes, programa matinal, jornal da tarde, telenovela de horário nobre e do fim da tarde e a programas do início da noite.

Para a TVI prevalecem os programas que correspondem a filmes ao longo de toda a semana, séries de ficção nacionais do horário nobre e do início noite, programas juvenis e talk-shows. Considerando as faixas etárias conforme podemos observar no círculo de correlações (*Gráfico F.5.5.III.3*) a faixa etária +64 anos é a variável que mais se distanciou da média, seguindo-se a adjacente (55-64) e os mais jovens (4-14). As restantes faixas etárias apresentam valores bastante correlacionados. Aos três primeiros eixos corresponde 88% de inércia explicada. O primeiro factor está mais correlacionado com as faixas 15-44 anos, o segundo com os mais jovens, o terceiro com 35-54 anos e o quarto com 25-54 anos. Portanto, com excepção do segundo factor abrangem várias faixas etárias adjacentes. Os mais idosos estão mais correlacionados com o segundo factor. Apenas os mais jovens apresentam um comportamento mais diferenciado.

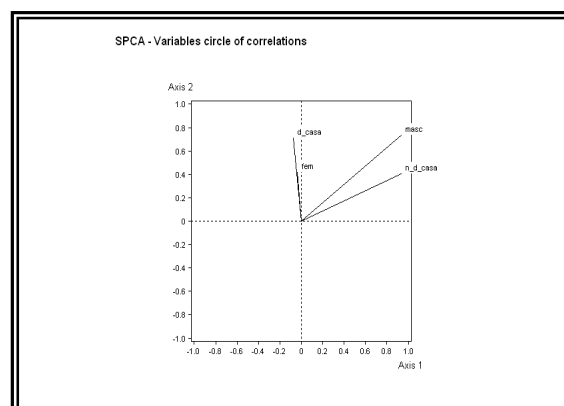


Gráfico F.5.5.III.4 - Correlação entre Índices de Afinidade para sexo e situação no lar

telenovelas do horário nobre, programas do Herman, entre outros. Para os segmentos sexo e ser responsável pelas compras do lar só é possível extrair um factor que apresenta maiores correlações e com valores positivos para o sexo masculino e para quem não é responsável pelas compras do lar e negativas para o sexo feminino e para as donas de casa (Gráfico F.5.5.III.4).

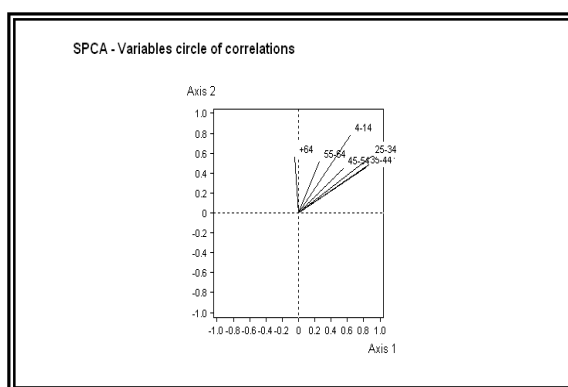


Gráfico F.5.5.III.3 - Correlação entre Índices de Afinidade para as faixas etárias

Podemos fazer as seguintes associações: O segundo factor está mais correlacionado com os programas juvenis, séries, filmes como os espaços RTP cinema (S/FS) noite, filmes Sábado e Segunda (TVI),

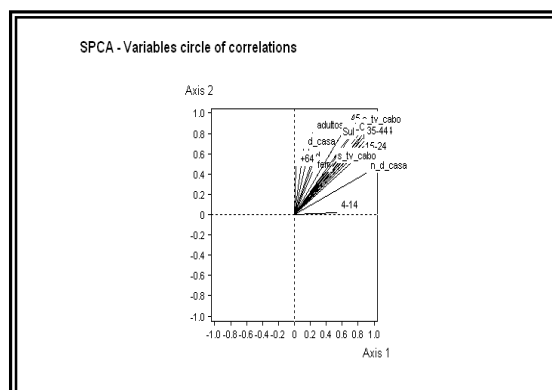


Gráfico F.5.5.III.5 - Correlação entre índices de afinidade para a totalidade dos segmentos

Ao primeiro factor associamos todas as regiões, excluindo o Sul e o Interior, todas as classes excluindo a classe D, masculino, 15-24, lares com TV cabo e telespectadores que não são donas de casa. Para o segundo factor associamos as mesmas regiões e o Sul, as mesmas classes sociais, o sexo masculino, 25-64, e também lares com TV cabo e as donas de casa.

Ao terceiro factor associamos todas as regiões, todas as classes excluindo a classe D, o sexo feminino, 15-54 e lares com TV Cabo.

Verificamos que os segmentos incluídos nas bases de dados, excluindo casos pontuais apresentam valores com uma ordem de grandeza semelhante. Estas variáveis de segmentação são demasiado abrangentes não nos permitindo a identificação de associações claras e diferenciadas entre programas e os segmentos.

Considerando os programas mais relevantes dos três canais em análise (espaços informativos, séries e telenovelas de prime time e noite), a aplicação da análise MDS simbólica, também, não produz bons resultados, com uma significativa concentração de SOs na origem do espaço MDS.

Conduzindo uma análise ACP simbólica o resultado gráfico também é pouco explícito.

Atendendo à globalidade da amostra, com base nas correlações das variáveis com os eixos e no contributo dos SO para os eixos, verificamos que o primeiro eixo factorial está correlacionado com as classes C1 e C2, com as idades compreendidas entre os 15 e os 44 anos, lares com TV cabo, quem não é dona de casa e com as regiões de Lisboa e Litoral Norte.

Neste apêndice coloca-se em evidência as potencialidades de determinadas áreas metodológicas quando se pretende realizar uma pré-selecção dos suportes a incluir nos planos de *media* e avaliou-se a capacidade discriminante das variáveis sociodemográficas ao nível das tipologias que maximizam a exposição televisiva para cada um dos canais generalistas.

Foi possível realizar algumas associações entre as variáveis sociodemográficas e os suportes, nomeadamente face às tipologias e canais, contudo estas variáveis condicionam a disponibilidade e o padrão de tempo de lazer ou vazio disponível, assim como a sua regularidade.

Logo, as associações podem decorrer deste factor e não obrigatoriamente de preferências ou da atractividade que determinados conteúdos podem suscitar nos telespectadores. Verificamos que em geral os índices sóciodemográficos apresentam uma considerável variabilidade se considerarmos suportes equivalentes no decorrer do tempo.

F.6. A EXTRACÇÃO DE REGULARIDADES ESTATÍSTICAS: AS ANÁLISES EXPLICATIVAS E DE PREVISÃO MULTIVARIADAS DE VALORES DE RATINGS PROGRAMÁTICOS

F.6.1. Introdução

Os modelos explicativos e de previsão de valores de *ratings* corresponde a uma temática que apenas pontualmente tem sido abordada na literatura internacional da especialidade e ainda em menor grau na bibliografia nacional. Os modelos explicativos são essenciais para que posteriormente possam ser desenvolvidas abordagens de previsão e que *Weber (2002)* também refere: “Which determinants are relevant for television audience prediction is one of the first questions that needs to be answered before work on the identification of a forecasting model.”

No *capítulo II* consta uma síntese dos reduzidos trabalhos internacionais que visam promover a previsão de indicadores de audimetria ou do número de telespectadores expostos a determinados programas e espaços comerciais, incluindo as abordagens qualitativas e as quantitativas. Neste capítulo privilegiam-se as análises quantitativas de Regressão Linear simbólica e as análises de Regressão não linear, assim como a RLM para efeito de comparação de resultados. As análises enquadram-se numa das áreas metodológicas que podem viabilizar a previsão das audiências, ou seja, a extracção de regularidades estatísticas.

Na área do *Data Mining* foi possível identificar o trabalho realizado por *Goodwin, Patelis e Assinkopoulo (2007)* que utilizam a Regressão Linear Múltipla, as Redes Neurais e os algoritmos de “Nearest Neighbours” para avaliar a repercussão da transmissão dos jogos de futebol nos *shares* de audiências, recorrendo a dados de audimetria gregos. Em *Weber (2002)* recorrendo a bases de dados de audimetria da Alemanha, a Redes Neurais e a algoritmos de *General Linear Models* tenta-se prever os valores dos *ratings* ao minuto, assim como uma variável binária que quantifica um acréscimo ou um decréscimo nos valores dos *ratings* quando os programas são alterados na grelha de programação.

Com a aplicação das Redes Neurais, *Weber* consegue valores de R^2 no intervalo [69,8%, 81,2%] para a previsão de valores de *ratings*. *Garland (1996)* com dados de audimetria Australianos recorre às Redes Neurais MLP para tentar prever os *ratings* ao nível do quarto de hora, desde a sexta até à novena semana de 1995, mas por análise das séries temporais univariadas.

As restantes publicações internacionais identificadas utilizam abordagens metodológicas que não estão associadas ao *Machine Learning* e nenhuma foi identificada na área da análise simbólica, existindo uma maior incidência em análises lineares que *a priori* não aparentam ser as mais adequadas. Por exemplo, em *Gensh e Shaman (1980)*, para a previsão de *ratings* os autores recorrem a Regressões Trigonométricas. *Henry e Heikki (1984)* para tentar prever os valores de *shares* recorrem a análises Multinomial Logit. De acordo com *Danaher, Dagger e Smith (2011)* a previsão de valores de *ratings* corresponde a uma área que tem sido negligenciada:

A review of the literature reveals that this is (the forecasting of ratings) a surprisingly neglected area of thorough academic and industry endeavour. ... The forecasting methods used in the past are not reliable and many have not been validated, also even more distressingly none have been tested in today's multichannel environment. ... Only one study used data from the 21st century. ... Earlier works include an overreliance on basic regression models and a lack of comparison of alternative forecasting models.

F.6.2. Objectivos

Antes do desenvolvimento de análises de previsão poderá ser necessário a identificação dos factores que condicionam em maior grau a evolução dos indicadores de audimetria e a avaliação das respectivas interdependências. Neste apêndice constam algumas análises multivariadas que essencialmente contribuem para o desenvolvimento de análises explicativas da evolução das taxas médias de audiências programáticas, uma vez que os instrumentos de previsão encontram-se condicionados pelos contextos em análise.

Segundo a bibliografia da especialidade, os modelos de previsão diferenciam-se em modelos explicativos ou univariados (Weber, 2002), mas um modelo multivariado apenas apresentará capacidade explicativa e predictiva se as variáveis identificadas forem conhecidas ou previsíveis quando se pretende realizar uma análise de previsão. As previsões que apenas recorrem à análise das séries temporais univariadas de valores de *ratings* apresentam uma menor relevância para a investigação televisiva e para a optimização do planeamento comercial, uma vez que não consideram factores “explicativos” ou que se associem com a sua evolução.

No entanto, no *Apêndice F.2* também constam algumas análises univariadas e que permitiram em alguns contextos alcançar razoáveis percentagens de variância explicada, constituindo uma satisfatória perspectiva de análise quando o seu principal objectivo reside apenas na previsão dos principais indicadores de audimetria.

Os denominados modelos explicativos das audiências incidem no desenvolvimento de análises matemáticas que viabilizem caracterizar a evolução dos índices de audiências, considerando variáveis que nem sempre são previsíveis, mas que condicionam a sua evolução, sendo esta a óptica privilegiada nesta tese.

Neste apêndice, pretendemos desenvolver modelos de caracterização da evolução dos índices de audiências e, não, de comportamentos individuais ou agregados de exposição televisiva. Inúmeros autores consideram que o consumo televisivo é aleatório e condicionado por inúmeras variáveis, muitas das quais impossíveis de quantificar o que poderá incutir uma maior variabilidade aos indicadores. Uma vez que as análises se debruçam sobre os dados de audimetria pretende-se avaliar até que ponto determinadas variáveis estruturais condicionam a evolução dos valores dos *ratings* programáticos globais e parciais, como, por exemplo, as variáveis associadas à caracterização dos programas, à estruturação da grelha, aos canais e temporais. Interessa avaliar a influência de determinados factores, para além das condicionantes temporais, ainda que este último factor aparentemente deverá ser o mais relevante inclusive indirectamente condicionando as interacções existentes entre as restantes variáveis consideradas nas análises.

F.6.3. Metodologias e bases de dados

Nas análises utilizam-se as bases de dados associadas ao *software* Telereport onde constam as variáveis de caracterização da grelha e os índices de audiências programáticos. Consideram-se apenas os valores associados ao universo televisivo e para a totalidade da emissão diária. Quando for necessário utilizar-se-á, também, uma base de dados referente a períodos horários, associando os valores das variáveis que constam das bases de dados Telereport a intervalos de tempo com uma duração de trinta minutos.

Nas análises a realizar considera-se um período característico a nível médio do consumo televisivo e idêntico aos das análises realizadas anteriormente. Todos os factores que condicionam a evolução da audiência média para o total da TV num dado momento apresentam uma inegável repercussão nos valores dos *ratings*.

Os dados de audimetria devem apresentar uma considerável variabilidade, o que se deve, quer à sazonalidade bem patente no consumo televisivo e às alterações nos padrões de disponibilidade da audiência, quer às frequentes alterações que se verificam na grelha de programas de cada canal e ao nível da contra-programação. O *rating* de um programa depende não só da faixa horária a que está associado, mas é, também, influenciado pela programação adjacente e concorrente. Os canais estruturam as grelhas de modo a maximizarem os fluxos de audiências entre os programas adjacentes. As estratégias de contra-programação adoptadas, também se diversificam no decorrer do dia, sendo por vezes não de confrontação directa, mas de complementaridade.

Como variáveis explicativas, que *a priori* parecem ser mais relevantes, vamos considerar o período horário em que o programa é transmitido e que resulta das análises de segmentação anteriormente realizadas, as tipologias dos programas e dos seus programas adjacentes uma vez que influencia o grau de duplicação de audiências; a caracterização da contra-programação pela sua repercussão ao nível dos *shares* de audiências e o canal em que o programa é transmitido, entre outras.

No entanto, para além destas variáveis inclui-se nas análises a totalidade dos índices que constam das bases de dados Telereport, assim como algumas variáveis de caracterização da grelha de programas.

Alguns dos índices mais relevantes constam do seguinte *Quadro F.6.3.1*. Ainda que a principal variável dependente em estudo corresponda às taxas médias de audiências (*Rat_Prog%*), também, é importante a análise das inter-relações existentes entre os restantes índices e as variáveis anteriormente mencionadas como sendo as mais relevantes do ponto de vista explicativo. Em especial a variável que quantifica a exposição global ao *media* e os *shares* também podem vir a ser objecto de análise.

Id	Descrição	Id	Descrição
Bsh%	Share de emissão de um programa. Corresponde ao rácio entre a duração do evento e a duração da emissão do canal naquele dia.	Ats#	Tempo médio despendido pela audiência. Corresponde ao tempo em minutos que cada telespectador despendeu em média a ver um canal ou evento.
Sat%	Índice de saturação. Este índice mede a proporção de inserções publicitárias na emissão de um programa.	Atv#	Tempo médio de visionamento. Corresponde ao tempo em minutos que cada indivíduo despendeu a ver um canal ou evento.
Rch%	Taxa de cobertura. Mede a % de indivíduos que contactaram com o suporte pelo menos 1 minuto. Se o mesmo indivíduo contactar duas ou mais vezes com o mesmo programa apenas é considerado um contacto.	Rsh%	Share de recepção de um programa. Mede a proporção de total de minutos contactados durante o programa no total de minutos contactados nesse dia no respectivo canal.
Shr%	Share de audiência de cada canal ou evento que é calculada a partir do tempo total despendido a ver esse canal ou evento relativamente ao tempo total despendido a ver televisão.	Alf%	Coeficiente Alfa que mede a relação entre o share de recepção de um programa e o seu share de emissão. Mede o contributo do programa para os níveis de audimetria do canal.

Quadro F.6.3.1 – Descrição de algumas das variáveis que constam das análises multivariadas a realizar
 Fonte: Elaborado pela autora.

Nas análises a realizar recorre-se às bases de dados com o formato clássico, ou seja, que cruzam variáveis e observações e onde consta apenas uma observação por célula, como também, às bases de dados simbólicas que representam dados agregados que caracterizam programas em concreto, espaços destinados à transmissão de filmes ou intervalos horários, tendo-se definido um Δt de 30, por exemplo, Espaço_Informativo_Noite_Fim_Semana ou 19H-19.30H. As variáveis numéricas clássicas transformam-se essencialmente em variáveis simbólicas da tipologia intervalo.

Para a realização das análises propostas recorre-se a algumas metodologias lineares e não lineares. Do ponto de vista linear implementaram-se metodologias de Estatística Multivariada e de análise Simbólica nomeadamente a Regressão Linear Múltipla.

As análises exploratórias traduziram-se no cálculo dos coeficientes de correlação linear de pearson e em medidas de associação entre as variáveis nominais e as variáveis dependentes em análise, quer dizer, em análises Anova, Testes T para valores Médios e na aplicação dos equivalentes testes não paramétricos. A análise factorial em Componentes Principais permitirá a avaliação da multicolineariedade existente entre as variáveis consideradas.

Do ponto de vista explicativo e predictivo as restantes abordagens permitem inter-relacionar as variáveis e descobrir as relações e o grau das inter-relações, assim como realizar uma previsão linear. As metodologias de Regressão Linear simples e múltipla e os modelos lineares por transformação dos dados iniciais são utilizadas para verificar o relacionamento funcional que, eventualmente, possa existir entre duas variáveis quantitativas. As análises de dependência são as que têm por objectivo o estudo da dependência de uma ou mais variáveis em relação às restantes, como, por exemplo, a análise Discriminante, a Correlação Canónica e a Regressão multivariada. Diferenciam-se das análises de séries temporais, em virtude destas últimas terem como objectivo modelar o estado do processo, extraindo e registrando desvios e tendências no decorrer do tempo.

A existência de interacções com um carácter não linear entre as variáveis explicativas implica a necessidade de recorrer às Árvores de Decisão, neste contexto de Regressão, uma vez que a variável critério é numérica. Nas análises a realizar implementaram-se os algoritmos CHAID e CART, os quais se diferenciam a diversos níveis.

O ajustamento dos modelos é quantificado nos modelos de regressão pela proporção de variância explicada. Um modelo não se deve induzir para a população de onde foi retirada a amostra, sem prevenir, o sobreajustamento. Consequentemente, sempre que necessário e atendendo à considerável dimensão da amostra, os dados iniciais foram subdivididos em amostra de treino, a qual serve de base para estimar o modelo explicativo e eventualmente também de previsão e em amostra de teste, em geral de menor dimensão, e sobre a qual se estimará o erro de previsão no final da construção do modelo. Alternativamente, recorre-se à validação cruzada, considerando dez subamostras.

Nas análises a realizar utilizaram-se, também, as Redes Neurais MLP–BP unidireccionais de aprendizagem supervisionada e as Redes RBF unidireccionais de tipo híbrido.

As limitações das Redes MLP traduzem-se na sua lenta convergência e em incorrerem facilmente no sobreajustamento aos dados. Consequentemente, para contornar esta última questão utilizou-se a validação cruzada, reservando um mínimo de 20% das observações para a amostra de teste e de acordo com o critério de que o número de observações para aprendizagem deverá ser $P = w \text{ (total de pesos)} / \varepsilon \text{ (erro de generalização)}$ para se conseguir um erro no valor de ε .

A amostra inicial subdividiu-se em amostra de treino que corresponde a 55% das observações, em amostra de validação (20% das observações) e em amostra de teste (25% das observações).

O algoritmo BP não garante que possa ser alcançado o mínimo global da função erro, apenas os mínimos locais. Logo, utilizaram-se diferentes matrizes iniciais de pesos, alterando a semente de geração de números aleatórios e utilizando as distribuições Uniforme e a Normal. No final selecciona-se a solução que proporcionar um menor erro de generalização.

Como funções de activação seleccionaram-se a tangente hiperbólica e a sigmoideia, representadas pelas seguintes funções:

$$f(x) = \frac{1}{1 + \exp(-bx)} \text{ (Sigmoidea)} \quad \text{e} \quad f(x) = b * (1 - f(x) * f(x)) \text{ (Tangente hiperbólica ou tanh)}$$

As funções tanh com valores no intervalo [-1,+1] aceleram a aprendizagem deste algoritmo. Para a função tanh, seleccionando a distribuição Uniforme e um intervalo de valores para os pesos dentro dos limites $[-2.4/\text{nin}, +2.4/\text{nin}]$ conseguimos uma convergência mais rápida do algoritmo, representando nin o número de entradas da Rede Neuronal.

No que diz respeito aos métodos utilizou-se o Gradiente Conjugado e o “Steepest Descendent”. Para o gradiente descendente a actualização dos pesos faz-se segundo a função

$$W_{(t+1)} = W_{(t)} + \alpha * (\partial E / \partial W) |_{W_{(t)}} + \beta * (W_{(t)} - W_{(t-1)}), \text{ W- pesos, } E(W) = S [E_i - V_i(W)]^2$$

Experimentaram-se diferentes configurações, variando as variáveis de entrada, o número de capas ocultas e o número de nós em cada capa oculta.

F.6.4. Resultados

I. Análises multivariadas de Regressão Linear

Neste apêndice pretende-se avaliar a capacidade explicativa e eventualmente de previsão de valores de *ratings* dos modelos construídos recorrendo à Regressão Linear Múltipla e apenas à informação que consta das bases de audimetria. Foram realizadas análises para a totalidade dos canais, assim como para cada um dos canais generalistas e considerando determinados períodos horários mais restritos. Em primeiro lugar avalia-se o grau de associação existente entre as variáveis explicativas e a variável dependente *Rat_Prog%* ou *Rat_Período%* se estiver respectivamente calculada para programas ou faixas horárias, considerando-se um Δt de 30 minutos. Posteriormente, constam os modelos de Regressão Linear Múltipla e de Regressão Linear Múltipla Simbólica por forma a comparar a capacidade explicativa e de previsão dos mesmos.

Ia. Análises exploratórias considerando a globalidade dos dados

Deste item constam os resultados das análises de correlação linear e de associação entre as variáveis nominais e *Rat_Prog%* ou *Rat_Período%*, recorrendo nas análises paramétricas à análise Anova a I Factor para as variáveis com mais de duas categorias e ao Teste T para comparação de médias entre amostras independentes para as variáveis dicotómicas. Nas análises não paramétricas recorreu-se ao Teste de Kruskal Wallis e ao Teste de Man Whitney U.

➤ Resultados dos coeficientes de correlação

A variável mais linearmente correlacionada com *Rat_Prog%* corresponde a *Alf_Prog* ($\rho=0,95$) que representa o contributo de um evento para os níveis de audiência desse canal.

Segue-se a taxa média de audiência para o total da televisão (*Tvr_Prog%*, $\rho=0.85$) e a taxa de cobertura representada pela variável *Rch_Prog* ($\rho=0,74$).

Apenas *Tvr_Prog%* e *Rch_Prog%* apresentam independência de cálculo face a *Rat_Prog%*. A variável *Tvr_Prog%* talvez seja mais facilmente prevista segundo as abordagens metodológicas sugeridas no *Capítulo VI*.

A taxa de cobertura mede a percentagem de indivíduos que contactaram com o programa pelo menos um segundo sem ponderar a duração de cada contacto e também está razoavelmente linearmente correlacionada com os *ratings*. A variável *Rch_Prog* diferencia-se de *Rat_Prog%* por não ter em consideração o tempo de contacto de cada telespectador com o suporte e se o indivíduo contactar duas ou mais vezes como o mesmo evento ou canal apenas é considerado como um contacto. Corresponde ao segundo indicador que também deverá ser alvo de análises de previsão univariadas, uma vez que poderá viabilizar a previsão dos *ratings* do ponto de vista linear. Alternativamente as abordagens de *Data Mining* não proposicional sugeridas no *Capítulo VI* também podem contribuir para a sua previsão.

Se considerarmos a correlação existente entre as variáveis *Rat_Período%* e *Tvr_Período%* ou seja, os valores associados a períodos horários em vez de a programas e para a totalidade da emissão diária, excluindo as observações atípicas, quer do próprio canal, quer dos canais concorrentes, o coeficiente R^2 é de 0,96 para o canal SIC, de 0,88 para a RTP1 e de 0,94 para o canal TVI. Se não excluirmos estes programas, como, por exemplo, os eventos desportivos os índices de correlação entre *Rat_Período%* e *Tvr_Período%* são de 0,93 para a SIC, 0,80 para a RTP1 e 0,91 para a TVI. Consequentemente, se não eliminarmos as observações atípicas penalizamos em maior grau a RTP1, sendo o decréscimo de apenas 0,03 nos coeficientes de correlação linear para os canais privados generalistas.

A RTP1 corresponde ao canal que apresenta em grande parte do período em análise um maior número de programas com observações atípicas. Portanto, ainda que não segmentando a emissão diária em faixas horárias os valores de correlação linear entre os valores de *ratings* e de taxas médias para o total da TV são muito elevados, especialmente se consideramos os dados referentes aos períodos horários, viabilizando a previsão dos *ratings* que posteriormente têm que ser extrapolados para os horizontes temporais que cada programa abrange.

No sector televisivo frequentemente considera-se como suporte um intervalo de 15 ou 30 minutos e os programas acabam por corresponder a um conjunto de suportes definidos a nível temporal.

Em vez de utilizar períodos de 30' podem ser considerados intervalos de tempo mais curtos, o que poderá levar a valores mais elevados do coeficiente de correlação.

A variável *Shr_Prog%* e *Shr_Período%*, ou seja, os *shares* de audiência referentes a programas e associados a períodos horários apresentam reduzidos coeficientes de correlação com os valores dos *ratings* para a totalidade dos canais. O coeficiente de correlação é de apenas 0,5 para *Shr_Prog%* considerando a totalidade dos dados e de 0,10 para o canal SIC, - 0,26 para o canal RTP1 e 0,48 para o canal TVI se considerarmos os valores de *Shr_Período%* e *Rat_Período%* excluindo os suportes com valores atípicos. Se não excluirmos estas observações os coeficientes são de 0,2 para o canal SIC, - 0,10 para a RTP1 e de 0,40 para o canal TVI. Consequentemente, do ponto de vista da Regressão linear se considerarmos a totalidade da emissão diária os valores de *shares* não se associam linearmente com os *ratings*.

O valor do *rating* e de *Tvr%* do programa posterior (*Rat_Ldout%*, $\rho=0,68$ e *Tvr_Ldout%*, $\rho=0,68$) também estão positivamente correlacionados com *Rat_Prog%*. Atendendo à globalidade dos dados as audiências de um programa estão mais linearmente correlacionadas com os valores do programa posterior do que com os *ratings* do programa anterior (*Rat_Ldin%*, $\rho=0,62$ e *Tvr_Ldin%*, $\rho=0,55$).

Com base nas análises realizadas aos efeitos de herança (Apêndice F.4) verificámos que a duplicação primária das audiências também é superior face aos programas sucessores e

avaliou-se a correspondência entre a evolução dos *ratings* e a duplicação primária das audiências, não existindo proporcionalidade entre os dois indicadores.

Se em substituição considerarmos os valores de *Ratings* e de *Tvr%* dos períodos adjacentes, os coeficientes de correlação são de 0,88 para *Rat_Período_Ldout* e de 0,88 para *Rat_Período_LdIn* para o canal SIC; 0,81 e 0,82 respectivamente para a RTP1 e de 0,93 e 0,95 para a TVI e não excluindo, programas mais populares.

Estes valores são muito superiores aos anteriores. Excluindo observações atípicas os valores de correlação são de 0,88 e 0,89 para a SIC, 0,85 e 0,86 para a RTP1 e 0,93 e 0,95 para a TVI. Estes valores são uma vez mais muito superiores aos que resultam de considerar programas. Verificamos a existência de linearidade entre os *ratings* de observações adjacentes considerando valores de períodos horários, não se diferenciando significativamente os valores de *lead_in* face aos valores de *lead_out*. Possivelmente, estes valores diferenciar-se-ão em função de períodos horários mais restritos.

As seguintes variáveis mais correlacionadas com os *ratings* são o *Atv* ($\rho=0,67$) e o *Rsh%* ($\rho=0,66$) as quais quantificam o tempo médio que cada indivíduo despendeu a ver um evento e o *share* de recepção de um programa.

A hora de início do programa (*Hora_Prog*) ou a sua conversão num número decimal (*Tempo_Time*) apresentam coeficientes de apenas 0,52 e a hora de conclusão de um programa (*Hora_Fim_Prog*) o valor de 0,46.

Considerando o factor temporal não existe linearidade na evolução dos valores dos *ratings*. Será necessário recorrer a horizontes temporais mais restritos, consequentemente posteriormente repetem-se as análises para os períodos horários que resultaram das análises de segmentação atendendo à evolução e variabilidade observada nos valores dos *ratings*.

As restantes variáveis apresentam valores dos coeficientes de determinação inferiores a 0,40. Os maiores coeficientes de correlação, mas negativos são irrisórios e correspondem a *Sat_Ldin%* (-0,15), *Ratio_Rch_Ldout* (-0,10) e *Dur_Espaço_Comercial_Ldin* (-0,10).

Uma maior saturação publicitária do programa anterior, o que obviamente está associado a uma maior duração de espaços comerciais no seu decorrer penalizam os *ratings* de um programa. Os programas com uma maior duração são transmitidos no período matinal ou no decorrer da tarde e correspondem aos períodos com menores níveis de audiências. O coeficiente com sinal negativo associado a uma maior cobertura de um programa face ao do programa seguinte é difícil de justificar, mas o valor é reduzido.

As restantes variáveis apresentam coeficientes irrisórios e dizem essencialmente respeito a durações dos programas e de espaços comerciais, saturações publicitárias, valores de fidelidade e sinergias entre tipologias de programas adjacentes. Portanto, do ponto de vista linear considerando a globalidade da amostra as durações ou a coincidência entre as tipologias não influenciam os *ratings*.

Com base nestes resultados verificamos que existem algumas limitações quanto ao desenvolvimento de modelos de Regressão Linear, pelos menos atendendo à totalidade da emissão diária e canais. Um reduzido número de variáveis de audimetria apresenta alguma correlação com a variável em estudo, a maioria das quais não são conhecidas no momento de realizar uma previsão.

Considerando períodos horários com uma evolução mais homogénea da variável em estudo, avaliar-se-á o grau de linearidade das relações e se existem alterações quanto ao grau de correlação entre as variáveis que constam das bases de dados de audimetria.

➤ **Avaliação do grau de associação existente entre determinadas variáveis nominais e *Rat_Prog%***

Realizando análises ANOVA e atendendo ao *factor canal*, a SIC apresenta uma média de 7,8% para *Rat_Prog%*, a TVI de 5,4% e a RTP1 de 4,8% (p valor =0,000). Era previsível que os canais privados apresentassem maiores valores de *ratings*.

Para as variáveis que traduzem a existência de compatibilidade de tipologias entre programas adjacentes, ou seja, as sinergias tipológicas de grau 1, 2 e 3 consoante nos referimos a determinado nível programático, não existe associação com *Rat_Prog%* (*Syn_Ldin1* - Teste T, p valor de 0,143; *Syn_Ldin2* - Man Whitney U, p valor de 0,237; *Syn_Ldin3* - Teste T, p valor de 0,55; *Syn_Ldout1*-Teste T, p valor de 0,24; *Syn_Ldout2*-Teste T, p valor de 0,70; *Syn_Ldout3* - Man Whitney U, p valor de 0,237).

No entanto, verificamos que as tipologias em 73% das observações não coincidem com a tipologia de Grau 1 do programa anterior, sendo estas percentagens de 84% para o Grau 2 e de 81% para o Grau 3. Quanto à coincidência com as tipologias dos programas sucessores estas percentagens são de 32% para o primeiro grau e de 62% para as restantes.

Entre as várias tipologias programáticas e *Rat_Prog%* obviamente que também existe alguma associação. Considerando as tipologias de Grau 1, os programas de Cultura e Juventude, essencialmente transmitidos no decorrer das manhãs ou no Canal 2 são os que apresentam menores médias de valores de *ratings* e as tipologias de desporto, ficção e informação são as que apresentam maiores valores (Tipologia de Grau 1- Anova, p valor=0,0000).

A tipologia de desporto está sem dúvida mais relacionada com o canal estatal e os programas lúdicos com os canais privados, para o período em análise. Os programas que não consistem em repetições apresentam maiores valores de *ratings* com um valor médio de 5,9% e de 4,6% caso contrário (Teste T, p valor=0,0043), assim como os que consistem em vários episódios face aos programas de uma única transmissão (Teste T, p valor=0,0000, média de 6,1% e 3,5% caso contrário).

Posteriormente, foram realizadas análises idênticas para cada canal. Com base nestes resultados foi possível concluir que todas as variáveis associadas a uma maior duração programática e a um maior índice de inserções publicitárias está associado a menores valores de *ratings* programáticos para o canal SIC e

para a RTP, mas para a TVI verifica-se o inverso. Obviamente que ambas as variáveis estão correlacionadas.

Um segundo aspecto a salientar é a relevância das variáveis associadas aos programas adjacentes. Para o canal SIC e a RTP1 os *ratings* dos programas sucessores apresentam maiores coeficientes de correlação linear face aos coeficientes dos programas anteriores e são essencialmente superiores para o canal estatal RTP1, uma vez que apresenta maiores desníveis nos valores dos *ratings* entre os programas adjacentes. Para a TVI estas variáveis equiparam-se e maximizam-se os seus valores.

Segundo alguns autores, existe uma tendência no público para assistir à televisão em sessões, até estarem satisfeitas determinadas necessidades, sejam elas de preenchimento de tempos livres, vazios, ou outras.

Consequentemente, antes e após assistirem aos programas ou faixas horárias pretendidas permanecerão em frente ao televisor por um certo período, variável em função de cada telespectador e no decorrer do tempo. Os programas que podem apresentar maiores associações com os valores de *lead in* correspondem essencialmente aos que seguem os espaços informativos ou os programas desde o início do período de prime time até às 22H-23H.

O canal estatal RTP1 é o que apresenta maiores valores de fidelidade e de *ratings* para os espaços informativos, mas, também, é o que favorece em menor grau os programas adjacentes.

No período da noite o programa que antecede o espaço informativo é favorecido, assim como o programa sucessor. A selecção do canal estará também condicionada pela notoriedade e imagem que o público lhes associa e que por sua vez gera comportamentos de lealdade programática a determinados espaços informativos.

A partir dos espaços informativos da hora de almoço verifica-se uma redução muito acentuada nos valores dos *ratings* até ao programa que antecede o espaço informativo da noite. Segundo estas análises os *ratings* programáticos estão mais linearmente correlacionados com a taxa média de audiência para o total de televisão e em muito menor grau com o valor do *share*, sendo este valor superior para RTP1, seguindo-se o Canal 2, a TVI e em último lugar a SIC.

Verificamos que até certo ponto a variável *Tvr_Prog%* é independente do conteúdo transmitido, sempre que não se verifiquem consideráveis alterações na grelha de programas.

A evolução desta variável, dia a dia, é muito semelhante, apresentando um primeiro pico de valores no espaço informativo do almoço e um segundo muito superior, uma vez que a população que exerce uma actividade laboral também estará disponível, em torno ao espaço informativo da noite que se prolonga pelo período de prime time e ao início da noite.

Com base nas análises de correlação verificamos que a variável *Tvr_Prog%* depende essencialmente de factores temporais

e que os seus valores estão linearmente correlacionados com os seus valores para os programas adjacentes e, também com os respectivos valores de *shares*.

Em relação aos valores de correlação entre *Tvr_Prog%* e os seus valores para os programas adjacentes, o valor sucessor é mais relevante para o canal estatal; o valor anterior destaca-se para a SIC e na TVI ambos os valores se equiparam. Portanto, para a RTP1 os programas saem beneficiados com a existência de uma maior audiência total disponível nos períodos seguintes e para a SIC verifica-se o inverso.

A variabilidade observada ao nível dos *ratings* associados a cada período horário é muito menor face à observada nos valores dos *shares*, logo a variável *Tvr_Prog%* será em parte responsável pelas fortes oscilações existentes neste último indicador.

Os intervalos de variabilidade de valores *shares* para os diversos canais generalistas são muito significativos, independentemente dos períodos considerados, o que também denota uma falta de lealdade programática e para com o próprio canal.

O facto da variabilidade observada ao nível dos programas ser muito superior à observada quando consideramos os períodos horários reforça esta ideia. As alterações que se verificam quanto aos horários de transmissão dos programas afectam substancialmente os valores dos seus *ratings*. Se entrarmos em períodos mais tardios os *ratings* decrescem de forma muito considerável quando se alteram os horários dos programas.

Portanto, a aparente passividade existente no consumo e a adesão do público aos constrangimentos estruturais verifica-se em maior grau entre programas directamente adjacentes decaindo fortemente com um maior distanciamento temporal entre os programas, uma vez que também impera a influência de um factor estrutural e que se traduz na disponibilidade de tempos vazios e/ou de lazer. Os efeitos de herança não são suficientemente fortes para manter um fluxo mais ou menos constante de telespectadores para a maioria das faixas horárias consideradas. Com base nas análises realizadas aos efeitos de herança do primeiro nível programático verificamos que estes coeficientes em geral não ultrapassam os 50% para programas directamente adjacentes, decaindo progressivamente com o distanciamento entre os programas.

Em relação aos resultados das análises ANOVA, segundo alguns autores a preferência por determinadas tipologias poderia favorecer os índices de audiência, desde que o planeamento favorecesse a transmissão contínua das tipologias mais apelativas à maioria do público e, assim, incentivar a maiores índices de duplicação de audiências entre programas adjacentes. No período de prime time transmite-se em primeiro lugar os programas informativos, seguindo-se os programas de ficção para os canais privados com a transmissão de séries, telenovelas, programas de humor e os concursos. A RTP1 é mais versátil transmitindo programas de humor, concursos, programas desportivos, filmes, programas de debate, entre outros.

Para o canal SIC e para a RTP1 os programas que não coincidem nas tipologias de grau 1 com os programas anteriores têm maiores valores de *ratings* (as restantes variáveis não são significativas) e para a TVI verifica-se o inverso, uma vez que os seus programas mais populares são as séries de ficção e as telenovelas nacionais. Interessa verificar em que medida os efeitos de herança estão apenas condicionados por factores temporais ou se também por parte das tipologias transmitidas e em função do canal. No *Apêndice F.4* constam algumas análises à repercussão de factores estruturais na evolução dos índices de duplicação primária.

Para todos os canais generalistas o facto de um programa não ser repetido ou estar associado a uma sequência de transmissões está associado a maiores valores de *ratings*.

➤ Avaliação da multicolineariedade entre as variáveis de audimetria

Uma vez avaliadas as correlações entre as várias variáveis explicativas e *Rat_Prog%*, *Rat_Periodo%* ou *Tvr_Prog%*, realizaram-se análises em Componentes Principais diferenciadas por canal e para a totalidade das observações para verificar o grau de multicolineariedade existente entre as variáveis e, assim, definir as abordagens de regressão mais adequadas.

Considerando os dados do canal RTP1 e incluindo a totalidade dos índices de audiências e variáveis de caracterização da grelha de programação, independentemente de se referirem ao programa ou à sua envolvente verificamos que um número significativo de variáveis estão correlacionadas, com ângulos

próximos a zero graus ($R^2=1$) ou a 180 graus ($R^2=-1$). Também, existem algumas variáveis independentes (ângulos de 90). As cinquenta e três variáveis iniciais quando reduzidas a cinco eixos factoriais obtemos uma percentagem de inércia explicada superior aos 70%. Através da análise dos coeficientes verificamos que o primeiro factor está correlacionado com as variáveis que em grande parte traduzem os índices associados ao programa anterior, como, por exemplo, *Rat_Ldin%*, *Tvr_Ldin%*, *Shr_Ldin%*, *Fid_Ldin%*, entre outras e como índices associados ao programa apenas constam as variáveis *Rat_Prog%* e *Tvr_Prog%*. O segundo factor associa-se a grande parte dos índices associados ao programa posterior e a alguns índices programáticos, por exemplo, *Rat_Prog%*, *Fid_Prog%*, *Shr_Prog%*, *Grp*, *Bet_Prog%*. Dada a correlação existente entre as variáveis recorre-se ao método *Stepwise* para o desenvolvimento de modelos de Regressão Linear.

Atendendo aos dados do canal SIC são necessários cinco eixos factoriais para ultrapassar os 70% de variância explicada. Para o canal TVI também são necessários seis eixos factoriais para atingir um mínimo de 70% de variância explicada. O primeiro factor associa-se aos índices mais relevantes, quer do programa como da envolvente, como, por exemplo, *Tvr_Prog%*, *Tvr_Ldin%*, *Tvr_Ldout* e o equivalente para *Rat_Prog%* e *Shr_Prog%*. Com base nestes resultados resulta evidente a existência de uma considerável correlação, entre algumas das variáveis, que varia de canal para canal. Consequentemente, primeiro aplicou-se a Regressão Linear Múltipla *Stepwise*.

II. Modelos de Regressão Linear Múltipla

As análises de Regressão Linear Stepwise foram realizadas para a totalidade dos dados e para alguns períodos horários mais restritos, que resultam da segmentação dos dados em função da variabilidade observada nos índices de audiências.

➤ Resultados para a globalidade dos dados e por canal generalista

Considerando a globalidade dos dados associados ao canal SIC, com apenas *Tvr_Prog%* e *Shr_Prog%* obtemos um coeficiente de determinação de 0,95 (*Beta Tvr_Prog%*=0,90 e *Beta Shr_Prog%*=0,36) e de 0.82 com apenas *Tvr_Prog%* (*beta*=0,82). Uma vez que os *ratings* resultam do produto das variáveis mencionadas um modelo que as inclua teria que apresentar um elevado coeficiente de determinação, sendo este também considerável quando incluímos apenas as taxas médias de audiências para o total da TV e apresentando-se esta última como a melhor alternativa para promover a previsão dos *ratings*. As restantes variáveis não apresentam uma grande relevância. Considerando a totalidade da amostra o período horário é a variável que condiciona em maior grau a evolução dos *ratings*. A dependência será razoavelmente modelada do ponto de vista linear considerando os valores de *Tvr_Prog%*, mas, não, linear no decorrer do dia considerando o factor temporal como variável explicativa. Consequentemente, será necessário restringir as análises a diferentes períodos horários mais restritos para avaliar o grau de dependência dos resultados face aos factores temporais e se outros factores

emergem como condicionantes da evolução dos *ratings*. Para o canal TVI assim como sucedeu com a SIC a variável *Tvr_Prog%* é a mais relevante apresentando um R^2 de 0.80. Se excluirmos estas variáveis os *ratings* adjacentes são as variáveis que assumem uma maior relevância, mas com um R^2 com um valor de apenas 0.60. As variáveis associadas aos factores temporais não apresentam coeficientes de determinação significativos. Para o canal RTP1 com a variável *Tempo_Time* dos programas adjacentes o R^2 é de apenas 0.33. Com a variável *Tvr_Prog%* o R^2 tem o valor de 0.83, sendo esta a variável que se associa linearmente com os *ratings*.

➤ Resultados considerando o horário nobre e por canal generalista

Considerando apenas os dados referentes ao período de prime-time do canal SIC, com *Shr_Prog%* (*beta*=1.02) obtemos um coeficiente de determinação de 0.82. Se incluirmos *Tvr_Prog%* ascende a 0.98. Os modelos incluindo as restantes variáveis apresentam valores pouco significativos do coeficiente de determinação. Logo, as análises lineares não são uma boa opção do ponto de vista da previsão. Se incluirmos apenas a variável *Tempo_Prog* o coeficiente R^2 é de 0.33. Logo, não existe uma evolução temporalmente linear nos *ratings* para este período, mas sem dúvida que quanto mais tardia for a sua conclusão mais penalizados são os programas. São essencialmente os valores dos *shares* dos programas que determinam os seus *ratings*, uma vez que considerando um horizonte temporal mais restrito existe uma maior estabilidade ao nível dos valores de *Tvr_Prog%*.

Considerando os dados do canal SIC desde as 19H que coincide com o início do pré-prime-time até ao final da emissão, a recta de regressão incluindo a variável *Shr_Prog%* apresenta um R^2 no valor de 0.51, ascendendo este valor a 0.97 se incluirmos *Tvr_Prog%*. Portanto, alargando o período horário os valores de *Tvr_Prog%* assumem uma maior relevância, uma vez que este período apresenta uma considerável tendência decrescente nestes valores.

Considerando as variáveis *Tempo_Time*, *Dur_Prog*, *Entre_Inicios_S* e *Tempo_Time_Out* o coeficiente é de apenas 0,40. Também, não existe uma evolução temporal do tipo linear, mas horários mais tardios de um programa a partir de determinado valor penaliza os valores dos *ratings*. Portanto, quando alargamos os períodos horários em análise a variável *Tvr_Prog%* será a que realmente condiciona os valores dos *ratings*. Por oposição quando restringimos as análises a faixas horárias mais homogéneas quanto à audiência total média disponível, a variável mais linearmente correlacionado com os *ratings* será o *share* de audiência (verifica-se para o prime-time).

Para o canal TVI com as variáveis *Tvr_Prog_Ldin%*, *Ratio_Rat_Ldin%* e *Shr_Ldin%* o R^2 é de 0.85. Os valores de *Tvr_Prog%* são uma vez mais muito relevantes, mas em especial o valor referente ao programa anterior. Com as variáveis *Tempo_Time_In*, *Dur_Ldin*, *Dur_Prog* o coeficiente de determinação é de 0.55. Portanto, para estes períodos não existe uma relação linear com o tempo, mas horários mais tardios e durações favorecem os *ratings*.

Também, é relevante o desempenho dos programas anteriores. Para o canal RTP1 com *Tempo_Time_In*, *Tempo_End* e *Dur_Prog* o R^2 é de 0.40 e com *Shr_Prog%* e *Tvr_Prog%* é de 0.98. Não existe uma associação linear com o tempo, mas quanto mais tarde for o início do programa menores são os valores dos *ratings*. Uma vez mais os valores dos *shares* são determinantes. Aparentemente para este canal os *ratings* são função em maior grau da atractividade que pontualmente determinados programas exercem do que por parte dos factores temporais. As audiências crescem linearmente ao longo da manhã para este canal e para o período em análise. A variável *Tvr_Prog%* neste contexto supera os valores dos *Shares*.

Nestes períodos e em especial para dias laborais uma vez que não existem alterações na programação e quase sempre também se mantêm os seus horários, o público apresenta uma certa lealdade ao seu canal e programa sendo os *ratings* essencialmente função da disponibilidade da audiência ou da predisposição para ligar o TV, a qual aumenta com uma maior proximidade face ao espaço informativo da tarde. Nestas faixas horárias também deverá existir uma maior incidência em telespectadores reformados e em maiores faixas etárias que são as que apresentam uma maior tendência a desenvolver uma maior fidelidade a determinados programas o que poderá contribuir para uma maior estabilidade nos seus níveis de audiências.

Para o canal TVI com *Tvr_Prog%* e *Shr_Prog%* obtemos um R^2 de 0.99 e com o *Tempo_Time* dos programas adjacentes o coeficiente é de 0.85.

Uma vez mais existe uma associação linear entre os *ratings* e os factores temporais no decorrer das manhãs pelas razões enumeradas anteriormente. Para a RTP1 com o *Tempo_End*, *Entre_Inícios (S)* e *Dur_Espaços_Comerciais_Antes* o coeficiente é de 0.71. Existe uma razoável associação linear com o tempo, assim como se verificou para os canais privados.

Considerando os dados do início da tarde do canal SIC com as variáveis *Tvr_Ldin%* ($R^2=0.80$), *Rat_Ldin%* e *Shr_Ldin%* obtemos um coeficiente de determinação de 0.98. O desempenho dos programas anteriores condiciona os *ratings* de um programa. Nestes períodos existe uma tendência descendente nos *ratings*, por conseguinte os maiores valores de *ratings* correspondem a programas que apresentam uma maior proximidade face a programas que apresentam maiores níveis de exposição ao painel.

Com as variáveis *Dur_Ldout*, *Tempo_Time_Out*, *Tempo_End* o R^2 é de 0.80. Quanto mais tarde finalizar o programa, o que também está associado a horas de início mais tardias dos programas seguintes e quanto mais longos forem os programas que sucedem menores são os valores dos *ratings*. Geralmente ao espaço informativo da tarde sucede um programa de menor duração e os *talk-shows* que apresentam várias partes estão associados aos declives existentes nos *ratings*. Portanto existe uma evolução consideravelmente linear neste período.

Para a TVI são mais relevantes os *ratings* dos programas adjacentes (*Rat_Ldout%* é a variável mais relevante) com um $R^2=0.80$.

Com as variáveis *Dur_Prog*, *Tempo_Time_In* e *Entre_Inícios_S* o coeficiente é de 0.55. Com uma maior duração programática e com uma hora de início mais tardia do programa anterior menores são os valores de *ratings*. Para a RTP1 com *Tempo_Time_Out*, *Dur_Ldout* e *Dur_Espaços_Comerciais_Antes* o coeficiente é de 0.52. Assim como, para a TVI apenas no período da manhã existe uma evolução linear.

Por último para o período da tarde e para a SIC com as variáveis *Tempo_Time_Out* ($R^2=0.72$), *Ratio_Tvr%_In* e *Rat_Ldout%* obtemos um R^2 de 0.96. Com as variáveis *Tempo_End* e *Dur_Ldin* o coeficiente é de 0.81. Também existe uma associação linear dos *ratings* com o tempo. Obviamente ao início da tarde os programas são influenciados pelos programas antecessores de maiores audiências da hora do almoço. No decorrer da tarde os programas mais beneficiados são os programas que estão mais próximo do início dos programas mais populares, ou seja, prevalece o efeito de *Lead_out*. No início da tarde existe um efeito de *Lead_in* e no decorrer da tarde de *Lead_out*. Considerando os dados do período da manhã do canal SIC com as variáveis *Tvr_Prog%* obtemos um coeficiente de determinação de 0.90. Considerando as variáveis *Tempo_Prog* e *Dur_Espaços_Comerciais_ntes*) o R^2 é de 0.77. Portanto, para este período a hora de início do programa está razoavelmente linearmente correlacionada com os *ratings*, favorecendo-os e a duração do programa que os antecede penaliza-os. Nos dois períodos existe uma relação linear com alguma variável temporal. Para a TVI a variável mais relevante é o *Shr_Prog%*, seguindo-se *Tvr_Prog%* com um $R^2=0.93$.

Para este canal apenas existe uma associação linear com o tempo no período da manhã. Uma vez mais *Tvr_Prog%* está linearmente correlacionada com os *ratings* para a maioria dos períodos.

III. Análise de Regressão Linear Múltipla Simbólica

Nas análises simbólicas os objectos de segunda ordem correspondem a determinados programas ou períodos horários ($\Delta t=30'$), caracterizados por variáveis da tipologia intervalo ou diagrama para além das variáveis métricas e nominais originais. Para a definição dos objectos simbólicos os programas foram agregados segundo o nome, canal e dia diferenciando a semana do fim-de-semana, por exemplo, os espaços informativos da noite de fim-de-semana de determinado canal, assim como todos os filmes do Sábado à tarde. Os espaços destinados à transmissão de filmes agregaram-se em função do canal, dia e hora de transmissão. Os modelos de Regressão podem incidir em programas em concreto e nesse caso a variável *Rat_Prog%* da base de dados Telereport será a variável critério ou em dados de *ratings* de períodos horários adoptando *Rat_Período%* como variável critério. Os períodos horários apresentam a vantagem de apresentar uma maior estabilidade de valores, desde que excluindo os eventos desportivos. Os programas apresentam uma maior variabilidade de valores essencialmente pelas alterações que se verificam nos seus horários e durações. A variável dependente nas análises a realizar corresponde a *Rat_Prog%* ou *Rat_Período%*, ambas da tipologia intervalo e as variáveis

explicativas coincidem com as que foram utilizadas nas análises de regressão multivariadas não simbólicas.

Como variáveis explicativas para realizar as análises de Regressão Linear simbólica atendendo aos programas em concreto dispomos de variáveis quantitativas que se transformam numa base de dados simbólica em variáveis da tipologia intervalo e as variáveis nominais mantêm-se ou são transformadas em variáveis diagrama ou taxonómicas. O canal, se o programa consiste em uma retransmissão, se é transmitido em episódios e as tipologias programáticas de grau 1 e 2 mantêm-se como variáveis nominais enquanto que o período horário, o dia, as sinergias de *Lead_in* e de *Lead_out* de ordem 1,2,3 correspondem a variáveis diagrama e as tipologias de grau 3 correspondem a variáveis taxonómicas.

As variáveis diagrama consistem numa distribuição de frequências para as várias categorias que a compõem. Agregando os dados que correspondem a determinado programa, obtemos um conjunto de períodos horários ou dias com os seus pesos correspondentes na amostra. Diferenciam-se os resultados em função dos canais e atendendo aos dados de períodos horários ou para programas em concreto.

Pretendeu-se verificar nomeadamente quais são as tipologias dos canais concorrentes que apresentam para cada canal generalista uma maior repercussão na evolução dos seus *ratings*. Avaliar em que medida a contra-programação realmente condiciona o desempenho de cada canal.

➤ **Resultados para o Canal SIC**

Atendendo aos dados do canal SIC referentes a períodos horários, ou seja, adoptando *Rat_Periodo%* como variável dependente foi possível verificar que um modelo de Regressão Linear Múltipla para a totalidade das observações e incluindo como variáveis explicativas os valores de *shares* dos vários canais, os *ratings* dos canais concorrentes e os valores de *Tvr_Periodo%* apresenta um R^2 de 0.95. As variáveis *Tvr_Periodo%* e *Shr_Periodo_SIC%* estão linearmente correlacionadas com *Rat_Periodo_SIC%*. Excluindo a variável *Tvr_Periodo%* são as variáveis *Rat_Periodo_TVI%* e *Rat_Periodo_RTP%* que apresentam coeficientes positivos e apenas *Shr_Prog_RTP1%* um coeficiente negativo. Os canais apresentam semelhanças quanto à evolução das suas taxas de audiências, correspondendo os maiores valores de *ratings* nos mesmos períodos para a totalidade dos canais, mas apenas maiores valores de *shares* para o canal estatal RTP1 penalizam o canal SIC em maior grau, para a maior parte do período em análise. A TVI para parte do período em análise apresenta uma menor repercussão nos valores dos *ratings* da SIC.

Considerando como variáveis explicativas os valores de *shares* e de *ratings* dos vários canais e o período horário, substituindo a variável *Tvr_Periodo%*, pela variável nominal período (períodos 1 a 6 que decorrem das análises de segmentação) e adicionando também as tipologias de grau 1, o factor período é o que apresenta coeficientes não nulos, com os períodos 5 e 3 a apresentar valores positivos e os restantes negativos ($R^2=0.80$).

Portanto, uma vez mais são os factores temporais que realmente influenciam a evolução dos *ratings*. Considerando a mesma base de dados, mas excluindo as observações atípicas, os *ratings* dos canais concorrentes estão positivamente correlacionados com *Rat_Periodo_SIC%* e neste contexto o valor associado à TVI assume uma maior relevância.

Os *shares* estão negativamente correlacionados com a variável em estudo e a TVI é o canal que penaliza em maior grau os *ratings* do canal SIC ($R^2=0.90$). Portanto, excluindo os eventos desportivos da RTP1 os dois canais privados apresentam evoluções de *ratings* mais semelhantes com a TVI obviamente a penalizar a SIC quando transmite programas de maior popularidade. A RTP apenas penaliza a SIC nos períodos associados ao desporto e estes foram excluídos da base de dados.

Adoptando *Shr_SIC%* como variável critério e analisando a influência das tipologias programáticas dos canais TVI e da RTP1 e do período (variável nominal) nos valores de *shares* da SIC esta é favorecida quando a TVI transmite musicais, humor, teatro no período de prime-time e é penalizada quando o canal RTP1 transmite Magazines desportivos, informação desportiva, musicais, noticiários, resumos de desporto, concursos, cerimónias de desporto, magazines desportivos, desenhos animados, documentários, acontecimentos especiais ou concursos juvenis.

As tipologias identificadas para a RTP1 correspondem essencialmente ao desporto e aos espaços informativos e seus programas adjacentes.

As tipologias da RTP1 que penalizam os shares da SIC são em maior número face às tipologias da TVI, uma vez que este canal apresenta uma programação mais semelhante. Ainda assim, sem dúvida que os factores temporais são a variável mais relevante para prever os ratings de qualquer canal.

Atendendo aos dados do canal SIC referentes a programas, ou seja, adoptando *Rat_Prog%* como variável dependente foi possível retirar as seguintes conclusões:

(1) Realizando uma análise de Regressão Linear Simbólica tendo como variável dependente *Rat_Prog%* (*T*) e como explicativas *Tvr_Prog%* (*T*), *Período* e *Tempo_Time*, a variável *Tvr_Prog%* é a que apresenta um coeficiente significativo com o valor de 0.29, ao qual corresponde um $R^2=0.87$. Sem dúvida que os intervalos de variabilidade das duas variáveis estão muito correlacionados.

Se incluirmos também o valor do *share* do programa o coeficiente associado a esta variável não é significativo. Os *shares* não se associam linearmente com os respectivos valores de *ratings* programáticos.

(2) Na ausência da variável *Tvr_Prog%* e incluindo os restantes índices referentes aos programas e aos seus programas adjacentes, em termos absolutos e relativos apenas os ratings dos programas adjacentes a nível relativo e absolutos estão linearmente correlacionados com os *ratings* de um programa.

A variável com maior peso corresponde ao *Ratio_Rat_Ldin%* (1.88), segue-lhe *Ratio_Rat_Ldout%* (0.77) e *Rat_Ldout%* e *Rat_Ldin%* com um peso idêntico (0.60). A este modelo corresponde um valor de R^2 de 0.81. Existe uma relação linear entre os *ratings* dos programas adjacentes .

Portanto, do ponto de vista linear é fundamental associar um intervalo de variabilidade a *Tvr_Prog%* ou *Tvr_Período%* para conseguir estimar os correspondentes valores de *ratings*. Os *ratings* programáticos também estão consideravelmente correlacionados com os valores dos programas adjacentes. Nos capítulos anteriores constatou-se que em alguns períodos esta associação afasta-se da linearidade e traduz-se essencialmente em evoluções polinomiais. O quociente entre o *rating* de um programa e o seu antecessor é a variável com maior beta.

➤ Resultados para o Canal TVI

Adoptou-se como variáveis explicativas o *período*, *Tvr_Período%*, as *tipologias*, os *ratings* e os *shares* e como variável critério o índice *Rat_Período%*. Considerando os dados referentes a períodos horários e atendendo à totalidade das observações se incluirmos a variável nominal *período horário* e o *rat_período%* e *shr_período%* dos canais concorrentes verificamos que os *ratings* da TVI estão positivamente correlacionados com os valores do canal SIC e negativamente com os do canal estatal.

Portanto, a TVI também poderá ser penalizada com os eventos desportivos do mesmo.

No entanto, é a variável período a que apresenta maiores coeficientes com o período 5 a ser o único com um factor de ponderação positivo ($R^2=0.91$). Se incluirmos a variável *Tvr_Período%* uma vez mais esta apresenta-se como determinante ($R^2=0.96$). Se incluirmos as *tipologias* dos três canais representada por uma variável modal e o *período horário*, os factores temporais são os que assumem coeficientes não nulos e as *tipologias* não são relevantes ($R^2=0.93$).

Adoptando como variáveis explicativas o *período*, *Tvr_Período%*, *tipologias*, *rating* e *shares* e como variável critério *Rat_Prog%* verificamos que os *ratings* do canal SIC e o *período horário* são as variáveis mais linearmente correlacionadas com os *ratings* do canal TVI ($R^2=0.93$). A SIC é o canal que apresenta um comportamento mais semelhante e uma vez mais os factores temporais são determinantes. Incluindo os *períodos* como variável nominal e as *tipologias* dos vários canais apenas a primeira é significativa. Se incluirmos *Tvr_Prog%* uma vez mais esta é a variável mais relevante.

Ao contrário do que sucedeu com a SIC os *ratings* dos programas adjacentes não estão linearmente correlacionados e os valores associados a *Ratio_Rat_In%* são inferiores à unidade para a maioria das observações, ou seja, os *ratings* dos programas são em geral inferiores aos dos programas anteriores.

➤ Resultados para o Canal RTP1

Atendendo aos dados de períodos horários e logo tendo como variável dependente *Rat_Período%* e incluindo

como variáveis explicativas *Tvr_Período%* e *Shr_Período* o R^2 é de 0.93. Excluindo estas variáveis *Shr_Período_SIC%* e *Shr_Período_TVI%* penalizam os *ratings* do canal estatal e está linearmente correlacionada com *Rat_TVI%* ($R^2=0.83$). Com apenas a variável *períodos* (nominal) e as *tipologias* da RTP1 os períodos apresentam coeficientes negativos com excepção do período 5 e apenas as *tipologias* de desporto favorecem significativamente os *ratings* ($R^2=0.94$). Portanto, para este canal os maiores valores de *ratings* associam-se aos espaços informativos em especial aos da noite e a alguns eventos desportivos.

Obviamente, que quanto maior forem os valores dos *shares* dos canais privados mais penalizado será este canal, o que se verifica para a maior parte da emissão diária sendo para este período mais penalizado por parte do canal SIC. Apenas existe uma maior correlação de valores de *ratings* com a TVI. Recorrendo a metodologias lineares verificamos que os resultados ao nível da previsão nem sempre são satisfatórios, independentemente dos horizontes temporais e canais considerados.

IV. Análises Multivariadas Não Lineares

Neste capítulo constam os resultados da aplicação de metodologias não lineares de Data Mining, nomeadamente das Árvores de Regressão e das Redes Neurais. A existência de interacções entre as variáveis explicativas e a não linearidade das mesmas torna relevante recorrer às Árvores de Decisão, neste contexto denominadas de Regressão uma vez que a variável critério é numérica.

Consideraram-se os algoritmos CHAID e CART. O algoritmo CART apresenta flexibilidade, mas sendo binário conduz a Árvores de grande dimensão e a resultados pouco eficientes e de menor interpretabilidade se a mesma variável tiver sido utilizada para subdividir os dados em vários níveis da Árvore.

Consequentemente, confrontaram-se os resultados com os que resultam da aplicação do algoritmo CHAID. Este algoritmo constrói Árvores com múltiplos nós a descender de cada nó que consta no nível anterior de desenvolvimento da Árvore. A sua principal diferença face ao CART traduz-se em evitar o crescimento da Árvore antes que ocorra o sobre ajustamento aos dados. O ajustamento dos modelos é quantificado nos modelos de regressão pela proporção da variância explicada.

Sempre que necessário e atendendo à considerável dimensão da amostra, esta foi subdivida em amostra de treino que serve de base para estimar o modelo de previsão e em uma amostra de teste em geral de menor dimensão e sobre a qual se estimará o erro de previsão no final da construção do modelo. Alternativamente, recorre-se à validação cruzada com $K=10$.

Nestas Árvores a previsão associada a cada nó terminal corresponde a uma constante que estima o valor da variável dependente Y . Supondo que o nó inclui os pontos (x_1, y_1) , (x_2, y_2) , ... , (x_n, y_n) a previsão corresponderá à média dos valores observados da variável dependente y_i para cada uma das observações incluídas nesse nó.

Considera-se a totalidade da emissão diária e as variáveis correspondem aos índices programáticos de primeiro e segundo grau e às variáveis de caracterização da grelha. A variável critério será $Rat\_Prog\%$ (T) quando se considera o primeiro nível programático e $Rat\_Prog\%$ (P) nas análises recorrendo aos dados do segundo nível programático. Realizam-se também algumas análises adoptando como variável critério $Shr\_Prog\%$.

➤ **Análises Recorrendo a Árvores de Regressão: Algoritmo CHAID e considerando a totalidade das observações**

A opção de validação utilizada nas análises foi a validação cruzada em que a amostra de teste corresponde a 30% do total da amostra. Da aplicação do algoritmo CHAID resulta uma variância explicada de 87%. A estimação do erro para a amostra de teste é de 2,20 e a variância total tem o valor de 17,07. As variáveis que mais se destacam correspondem a $Tvr\_Prog\%$, *ratings* dos programas adjacentes, tipologias e canais. Segundo a Árvore que maximiza a variância explicada, a variável mais relevante e utilizada no primeiro nível é $Tvr\_Prog\%$. As taxas médias de audiências são essencialmente função das taxas médias de audiências para o total da TV, uma vez que esta variável apresenta uma grande variabilidade no decorrer de um dia e acaba por condicionar as variáveis que mais se associam à variável em estudo. O *rating* de um programa é função do período horário a que está associado, uma vez que este influencia as estratégias de contra programação adoptadas; a estruturação da grelha de programas, a composição sociodemográfica da audiência e

determinados factores estruturais, como por exemplo, o grau de duplicação de audiências entre os programas adjacentes.

Posteriormente as variáveis *Rat_Ldin%*, *Rat_Ldout%*, as *Tipologias programáticas* e o *Canal* são as variáveis mais utilizadas no desenvolvimento da Árvore. O *rating* de um programa será função do grau de exposição à programação adjacente, da notoriedade do canal a que está associado e das tipologias programáticas envolvidas. Considerando apenas estas variáveis conseguimos uma razoável percentagem de variância explicada, mas que do ponto de vista predictivo fica aquém, uma vez que os *ratings* adjacentes não são facilmente previsíveis. No *Apêndice F.4* também se demonstrou que a ordem de grandeza dos *ratings* adjacentes não reflecte os valores de duplicação das audiências.

Verificamos que quanto mais elevado for o valor de *Tvr_Prog%* maior será a associação entre o *rating* de um programa e os *ratings* dos programas adjacentes e em especial para os canais privados generalistas, onde também mais facilmente podem ocorrer os efeitos de herança para o período em análise. Em função dos valores de *Tvr_Prog%* poderá prevalecer uma maior associação com o *rating* do programa anterior ou sucessor.

Atendendo ao período em análise, para os valores mais elevados de *Tvr_Prog%* e superiores aos 28% ou no intervalo [15%, 22%] a variável *Rat_Ldin%* supera *Rat_Ldout%*, ou seja, existe uma maior associação com a audiência do programa anterior e em menor grau com o índice posterior.

Correspondem essencialmente a programas de prime-time ou da noite, repartidos por todos os canais e que numa elevada percentagem estão associados aos espaços informativos. Este resultado poderá contribuir para confirmar que os programas adjacentes aos espaços informativos, em especial os subsequentes saem beneficiados. Os espaços informativos funcionam como pilares da grelha de programas. Os efeitos de herança em sentido inverso são mais reduzidos, ou seja, é maior a duplicação de audiências entre um programa e o seu sucessor directo do que perante o programa anterior. Estes resultados acompanham a evolução e associação entre os *ratings* adjacentes, uma vez que verificamos que é maior a associação entre *Rat_Prog%* e o valor de *Rat_Ldin%*.

Além do mais quanto mais nos afastamos dos espaços informativos após um pequeno período de maior estabilidade nas audiências tem início uma tendência decrescente, afastando cada vez mais *Rat_Prog%* dos *ratings* associados aos programas posteriores. Obviamente que no período com início no espaço informativo será onde existe uma maior fracção da audiência disponível por um determinado período, sendo este mais alargado para os canais privados do que para a RTP1.

Consequentemente, a duplicação de audiências poderá ser menos relevante para a RTP1 onde existe uma selecção programática mais pontual e em menor grau influenciada por condicionantes estruturais.

Para os valores de *Tvr_Prog%* que correspondem essencialmente a programas no decorrer de toda a *tarde, manhã e noite*, repartidos por *todos os canais*, prevalece a associação com *Rat_Ldout%*. Correspondem aos períodos adjacentes aos períodos de maior audiência e a sua proximidade face ao início de programas que apresentam uma maior exposição favorece-os, logo *Rat_Ldout%* será a variável mais relevante neste contexto.

Segundo as análises exploratórias, a tarde encontra-se subdividida em dois períodos com tendências opostas, ou seja, crescente e decrescente; a manhã apresenta uma evolução crescente e a noite uma evolução decrescente de valores de *ratings*. Obviamente que *Rat_Ldout%* assume uma maior relevância nos períodos crescentes de valores de *ratings*, mas também não deixa de ser válido que em períodos de tendência descendente se forem colocados programas mais aliciantes em períodos um pouco mais tardios isto poderá contribuir para que se observem valores de *ratings* mais elevados no seu decorrer dada a expectativa do público pelo início do programa.

Para os *canais estatais* e reduzidos valores de *Tvr_Prog%* (valores <6,6%) o que mais influência os *ratings* será a *tipologia* do programa. Os telespectadores assistem a programas específicos, como, por exemplo, aos espaços informativos, não se observando de forma tão significativa os efeitos de herança entre os programas.

Por exemplo, os jogos de futebol transmitidos neste canal apresentam uma grande exposição, mas por vezes não influenciam fortemente as audiências dos programas adjacentes. A excepção corresponde ao espaço informativo quando este é transmitido após o evento desportivo. Nas situações em que o seu horário é antecipado o espaço informativo fica penalizado como consequência dos padrões de disponibilidade da audiência e de selecção do media e possivelmente, também por desconhecimento da programação.

Os programas inseridos no intervalo de valores de *Tvr_Prog%* mais reduzido (<6,6%) correspondem em 40% ao final da noite, 48% ao período da manhã, 6,5% ao final da tarde e 4,6% ao princípio da tarde; 20% ao Sábado, 17% ao Domingo e os restantes dias em torno aos 10%.

Portanto, estão essencialmente associados à manhã, tarde e ao final da noite nos quais a tendência para os efeitos de herança é menor. Verifica-se uma selecção mais pontual que será essencialmente função da tipologia do programa. As tipologias que apresentam menores valores de *ratings* correspondem essencialmente aos filmes de fim-de-semana.

Consequentemente, segundo este algoritmo os *ratings* dependem significativamente das tipologias, desempenho dos programas adjacentes, canais e em muito maior grau dos factores temporais.

➤ **Resultados segundo o algoritmo CART para a Totalidade das Observações (SPSS)**

Da aplicação do algoritmo CART resultou uma variância explicada de 82%, com base na amostra de teste, sendo a variância total de 17,07; a Estimação de risco 3,14 e Et da estimação de risco de 0,63. As variáveis que se destacam coincidem com as do algoritmo CHAID, mas supera ligeiramente a percentagem de variância explicada.

A variável mais relevante e utilizada no primeiro nível da Árvore, corresponde uma vez mais a *Tvr_Prog%*. Para os valores de *Tvr_Prog%* mais elevados, os programas beneficiam de valores de *Rat_Ldin%* e de *Rat_Ldout%* inseridos também nos maiores intervalos o que se traduz em uma média de 11.9% para *Rat_Prog%*. Correspondem a programas inseridos entre programas que também apresentam maiores valores de *ratings* durante o prime-time e o início da noite onde os *ratings* se mantêm em níveis mais elevados por um maior horizonte temporal.

Para os mesmos níveis de *Tvr_Prog%*, mas para menores valores de *Rat_Ldout%* e que se traduz numa média de 7.9% para *Rat_Prog%* os programas correspondem ao final da noite essencialmente associados às tipologias de concursos, telenovelas, galas, talk-shows e desporto. Para valores de *Rat_Ldin%* num intervalo inferior as categorias de concursos, telenovelas, humor, telejornais, séries e desporto apresentam maiores valores de *ratings*.

Para valores de *Tvr_Prog%* no intervalo intermédio, as tipologias de magazine,

concursos, telenovelas, telejornais, desporto e talk-shows com maiores valores de *Rat_Ldin%* associados, maximizam os valores de *ratings*. Para as restantes tipologias programáticas de grau 2 os programas do fim-de-semana e dos canais privados apresentam maiores valores de *ratings*. Para os programas transmitidos de Segunda a Sexta, as tipologias de filmes e informação desportiva apresentam maiores *ratings*.

Para valores de *Tvr_Prog%* no menor intervalo de valores os programas que lhe estão associados, obviamente que minimizam os valores dos *ratings*.

Estes resultados diferenciam-se dos associados ao algoritmo CHAID por a Árvore apresentar apenas três intervalos de valores de *Tvr_Prog%*. A maiores valores desta variável corresponde uma maior associação entre o rating de um programa e os *ratings* dos programas adjacentes. Para maiores valores de lead in e menores de lead out saem favorecidos as tipologias de concursos (por ex. concurso da RTP1, após o qual se verifica uma forte redução nas audiências), as telenovelas de prime time antecederidas pelos espaços informativos que apresentam grandes audiências e após as quais se dá início à tendência decrescente, o desporto (programas essencialmente de RTP1), as galas e os talk shows. Aos valores de *Tvr_Prog%* no intervalo intermédio correspondem maiores valores de *ratings* às tipologias de magazine, concursos, telenovelas, telejornais, desporto e talk-shows. Os programas com os menores valores de *Tvr_Prog%* também correspondem aos que minimizam os *ratings*.

Os resultados produzidos pelo algoritmo CHAID são mais facilmente interpretáveis e lógicos. O algoritmo CART não produziu resultados muito satisfatórios, uma vez que utiliza partições binárias e recorre às mesmas variáveis em diferentes níveis do desenvolvimento da Árvore. Consequentemente, confrontam-se estes resultados com os que decorrem da aplicação dos algoritmos do *software* SAS onde é possível aproximar os mesmos às características do CART. Diferenciam-se os resultados para a totalidade das observações assim como para determinados canais e/ou períodos horários.

➤ Resultados atendendo à totalidade das observações e recorrendo ao algoritmo CART (SAS)

A variável utilizada no primeiro nível da Árvore corresponde a *Rat_Ldin%*. A percentagem de variância explicada é de 85%. Segundo este algoritmo as taxas médias de audiências associam-se em maior grau com o rating do programa que lhes antecede e não com a variável *Tvr_Prog%*. Verifica-se uma maior proximidade de valores de *ratings* face ao valor interior. Uma vez mais os *ratings* apresentam uma maior associação com *Rat_Ldin%*, uma vez que os efeitos de herança processam-se essencialmente em sentido directo, apesar da relação entre ambas as grandezas não ser necessariamente proporcional.

Para valores de *Rat_Ldin%* inseridos no maior intervalo de valores a influência advém dos períodos horários em que os programas estão inseridos.

Nos períodos de maiores audiências, nomeadamente no pré-prime-time, prime time e ao almoço, as tipologias (para programas muito populares) e os níveis de *Tvr_Prog%* influenciam os resultados. Os programas desportivos maximizam as audiências e para as restantes tipologias os *ratings* são essencialmente função dos níveis de *Tvr_Prog%*. Sem dúvida que os eventos desportivos podem maximizar os níveis de *Tvr_Prog%* e dos *Ratings* (19,6% de média de *Rat_Prog%*) diferenciando-se quanto ao grau de repercussão nos programas sucessores e concorrentes. As restantes tipologias não influenciam significativamente os níveis de audiências da contra programação a avaliar pela existência de uma certa estabilidade na ordem de grandeza dos *ratings* ao longo das semanas, cuja variabilidade atribui-se essencialmente aos factores de disponibilidade e de predisposição para ver TV do painel. Para estes programas são os níveis de *Tvr_Prog%* que mais os influenciam.

Se o intervalo de valores de *Tvr_Prog%* for o maior então a média de *Rat_Prog%* será de 10,1% e 7,2% caso contrário. Verificámos que para estes períodos existem oscilações no valor de *Tvr_Prog%*, estando os seus valores máximos associados ao prime-time. No pré-prime-time existe uma tendência ascendente nestes valores e a hora do almoço corresponde ao segundo pico diário de audiências ainda que inferior ao da noite. Nos períodos de menores audiências e que também abrange o final da noite e a noite os programas beneficiam de maior proximidade face aos programas mais populares e em especial com os que lhe antecedem (valores de *Rat_Ldin%*).

Por exemplo, os programas que sucedem aos espaços informativos ou a telenovelas. Para valores de *Rat_Ldin%* inferiores ao segundo intervalo de valores, os *ratings* estão mais associados aos valores posteriores, ou seja, a *Rat_Ldout%* e obviamente quanto mais elevados estes forem mais beneficiam os programas, assim como os valores de *Tvr_Prog%*. Isto sucede, por exemplo, em períodos com tendências ascendentes nos dados. Quanto maior forem as audiências posteriores maior será o valor de *Rat_Prog%*, o que também corresponde às horas mais tardias dos períodos referidos e a maiores valores das taxas médias de audiências para o total da TV.

Estes resultados estão influenciados pelas tendências anteriormente apontadas para o canal estatal RTP1 e pela presença dos jogos de futebol. Se excluirmos tais observações atípicas, mas considerando a totalidade dos canais, uma vez mais *Rat_Ldin%* é a variável utilizada no primeiro nível.

Para valores de *Rat_Ldin%* no menor intervalo de valores, *Rat_Prog%* está condicionado por *Rat_Ldout%*. Em períodos de menores audiências globais, os maiores valores de *Rat_Ldout%* favorecem os programas. Este impacto também depende dos níveis de *Tvr_Prog%* e dos *canais* considerados, sendo minimizados para o *Canal 2*.

Se *Rat_Ldin%* pertencer ao segundo menor intervalo, tudo dependerá do período horário em questão e dentro de cada período uma vez que as variáveis categóricas abrangem um conjunto de suportes horários, dos níveis de *Tvr_Prog%*.

Verificamos que os níveis de audiência média para o total da TV nestes períodos pode diferir significativamente. Obviamente que quanto maior for *Tvr_Prog%*, também maiores serão os valores dos *ratings*. Posteriormente, para maiores níveis de *Tvr_Prog%* as audiências dependem de *Rat_Ldin%*. Os maiores valores de ratings estão próximo dos programas mais populares.

Considerando a totalidade das observações são os factores temporais que influenciam os resultados e dada a evolução gradual existente nas audiências também os valores adjacentes que em função dos períodos alternam em relevância e quanto a uma maior similitude com a envolvente. Verificámos que os índices de duplicação entre programas directamente adjacentes podem ser razoáveis e a existência de proximidade entre os valores de *ratings*. Obviamente que em períodos de tendências ascendentes a incorporação gradual de novos elementos do painel é o factor responsável por esta evolução uma vez que os índices de duplicação são sempre inferiores à unidade.

➤ Resultados atendendo aos dados referentes apenas aos canais privados e recorrendo ao algoritmo CART (SAS)

Interessa também restringirmos as análises aos canais privados, uma vez que os canais estatais apresentam uma evolução mais atípica e que condiciona fortemente os resultados anteriores. Considerando apenas os canais privados SIC e TVI com apenas dois níveis na Árvore de Regressão conseguimos uma variância explicada de 85% e também por aplicação do algoritmo CART.

Os *ratings* programáticos estão condicionados em primeiro lugar pelos níveis de *Tvr_Prog%*. Portanto, concluímos que para os canais privados são os factores de disponibilidade e de vontade de ver TV e os horários em que os programas são transmitidos os que mais influenciam os *ratings*. Para os valores mais reduzidos de *Tvr_Prog%* os valores são apenas função do valor deste indicador e para os valores mais elevados uma vez mais as audiências estão também condicionadas por variáveis associadas a factores temporais, como, por exemplo, o tempo de conclusão de um programa e o período horário.

Para valores mais elevados de *Tvr_Prog%* os programas com horários de conclusão após as 19.30H apresentam maiores índices audiências e destes os programas associados ao prime time. Para horários de conclusão anteriores às 19.30H, mas com valores de *Tvr_Prog%* > 20.25% as maiores audiências também estão associados a maiores valores de *Rat_Ldin%*.

Se excluirmos das análises os valores dos *ratings* adjacentes e que corresponde a uma Árvore com 90% de variância explicada para os canais privados, verificamos que os níveis de *Tvr_Prog%* condicionam a associação existente entre as restantes variáveis explicativas e *Rat_Prog%*. Para níveis de *Tvr_Prog%* > 20% e logo mais elevados e com *Tempo_End* < 0.88, portanto programas que finalizam antes das 19.30H, os maiores valores de *ratings* estão associados a maiores valores de *Fid_Ldout* (se *Fid_Ldout* > 38,55% *Rat_Prog%*_Médio é de 10% e 7,87% caso contrário).

O que influencia os *ratings* é a fidelidade da audiência aos programas posteriores. Por exemplo, verifica-se à hora do almoço, com uma maior proximidade face ao espaço informativo a beneficiar os programas antecedentes.

O período designado de prime time desagrega-se nos programas transmitidos durante a semana ou de fim de semana, apresentando estes últimos menor média de *Rat_Prog%* e no valor de 11,2% face aos 13,7% para período semanal.

Para os períodos de maiores audiências os *ratings* são influenciados uma vez mais por factores temporais com a semana a sobressair uma vez que em dias laborais será menor a iniciativa do público para recorrer a outras actividades lúdicas e prevalece a necessidade de preenchimento de tempos vazios. Ao fim-de-semana dependendo das condições climáticas algumas pessoas recorrem a outras formas de preenchimento dos seus tempos livres penalizando as audiências televisivas.

Para os períodos de pré-prime-time e noite, os programas com maiores valores de *ratings* podem ser diferenciados pela fidelidade aos programas posteriores ou pelo canal. Uma maior fidelidade aos programas que lhe sucedem beneficiam os programas.

O conteúdo programático pode ter alguma repercussão nas audiências.

Também, foi possível verificar com base nas análises exploratórias realizadas aos fluxos de audiências que existe alguma intensa actividade dos telespectadores no decorrer de um programa investigando as alternativas existentes, assim como entre os programas transitando entre eles. Contudo, outras variáveis acabam por condicionar em maior grau a evolução dos *ratings*.

Em relação ao segundo subconjunto (Fig. F.6.4.1) de programas diferenciado no primeiro nível da Árvore (*Tvr_Prog%*<20%), para além dos níveis de *Tvr_Prog%* condicionarem fortemente as audiências, para valores de *Tvr_Prog%* superiores aos 15% surgem as tipologias de grau 2, como sendo a variável que condiciona em maior grau os *ratings*. Alternativamente, poder-se-ia utilizar a variável *canal* ou *Fid_Ldin*.

As tipologias mais favorecidas são as telenovelas, a informação especializada e os talk-shows e em especial os programas da TVI (valores de *Rat_Prog%*=8,34%).

Para as restantes tipologias é o período horário que as diferencia, sendo os *ratings* maximizados para os períodos do almoço, início da tarde, noite e final noite (6,1%, 4,82% caso contrário).

Alternativamente, em vez de recorrer à variável tipologia de grau 2, a SIC corresponde ao canal mais favorecido e os programas antecidos por outros aos quais o público lhes dirige maior fidelidade.

Portanto, para os menores valores de *Tvr_Prog%* os programas diferenciam-se com base nos níveis desta variável e apenas para valores superiores aos 15% surgem as tipologias programáticas. Existem maiores valores de *ratings* associados a determinadas telenovelas e talk-shows da TVI. Para as restantes tipologias é o período que as condiciona ou o canal e a fidelidade dirigida aos programas anteriores.

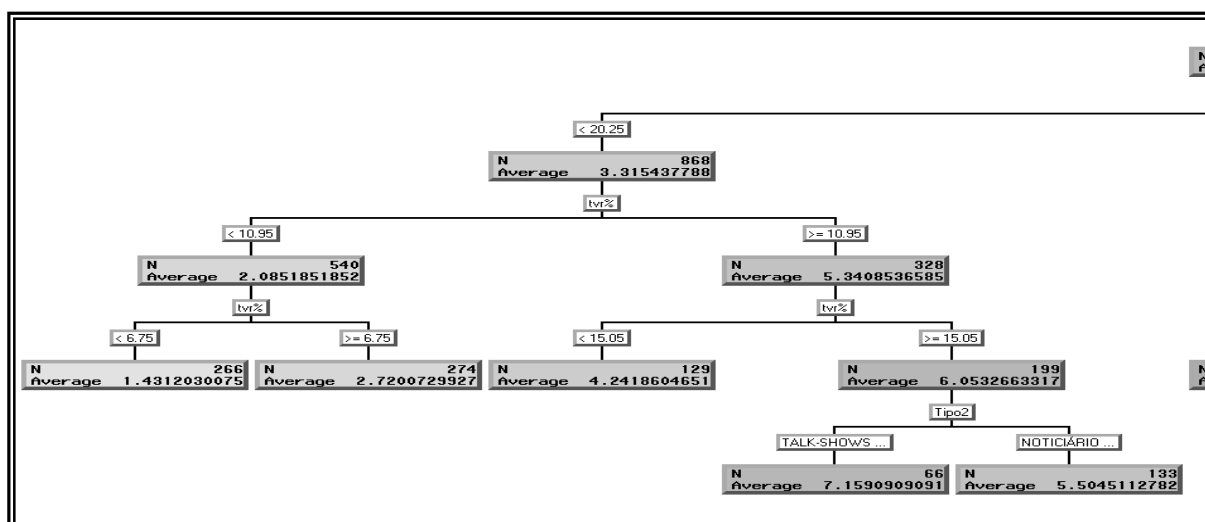


Fig. F.6.4.1 – Representação da Árvore de Regressão CART

➤ **Resultados com dados referentes aos canais privados e restringindo as análises aos períodos de pré-prime-time, prime-time e noite (SAS).**

A Árvore resultante apresenta 80% de variância explicada com uma amostra de teste corresponde a 30% das observações. A hora de conclusão dos programas é a variável que influencia primordialmente estes resultados, sendo utilizada no primeiro nível da Árvore.

Para um horário de conclusão dos programas anterior às 21.30H, os níveis de *Tvr_Prog%* influenciam os valores dos *ratings*, surgindo posteriormente *Rat_Ldin%* no desenvolvimento da Árvore. A este nó corresponde 8,33% de valor médio de *Rat_Prog%* e 12,3% caso contrário.

No decorrer do período de pré-prime-time, existindo uma tendência ascendente nas audiências, obviamente os programas serão beneficiados com a existência de maiores audiências para o total da TV e com a popularidade dos programas antecessores.

Em especial para o canal SIC os *ratings* são superiores após o espaço informativo da noite. Consequentemente, para horários de conclusão após as 21.30H os maiores valores de *ratings* estão associados a maiores valores de fidelidade (se *Fid_Prog%*>44% então *Rat_Prog_Médio%*=13% e 10,7% caso contrário). Para os maiores valores de fidelidade destacam-se as telenovelas, as séries, o humor e passatempos.

Para índices de fidelidade inferiores aos 44%, os programas são favorecidos por maiores valores de *Rat_Ldout%*, portanto saem

beneficiados com a proximidade a programas aos quais existe uma maior exposição. Por isso, mesmo os canais, por vezes adoptam como estratégias de programação a transmissão de programas menos apelativos entre programas de maior notoriedade.

➤ **Resultados com Dados referentes ao canal estatal RTP1 (CART - SAS)**

Uma vez concluídas as análises aos canais privados se restringirmos as análises ao canal RTP1 a variável utilizada no primeiro nível (Árvore com 80% de variância explicada) é o *período horário*, sendo os valores de *ratings* maximizados à hora do almoço, pré-prime-time e prime-time (9,23% de média de *Rat_Prog%* face aos 2,3% para os restantes períodos).

Nestas faixas horárias os níveis de *Tvr_Prog%* condicionam os *ratings*. Quando estes valores são muito elevados (*Tvr_Prog%*>41,15%), o que corresponderá a determinados jogos de futebol atingimos os maiores valores de índices de audiências (22,96% de média de *ratings* e 8,91% caso contrário). Neste contexto a *hora de conclusão* do programa condiciona os valores (10% de média de *Rat_Prog%* e 7,2% caso contrário) e se esta for após as 20H quanto maior for o valor do *share* maiores serão os valores de *ratings* (*Shr_Prog%*>26% corresponde a *Rat_Prog_Medio%* =11,23% e 8,5% caso contrário). Portanto, os *ratings* são maximizados pelos eventos desportivos e os restantes programas dos períodos de maiores audiências após as 20H quanto mais conseguirem maximizar o seu *share* de audiências.

Isto confirma a selecção mais pontual que se verifica neste canal e que não está condicionada à popularidade dos programas adjacentes. Verificamos uma maior relevância dos factores de planeamento, nomeadamente dos factores temporais de início e conclusão.

Para o pré-prime-time ($Hora\_End < 20h$) são os níveis de $Tvr\_Prog\%$ que influenciam os ratings (se $Tvr\_Prog\% > 17,1\%$, $Rat\_Prog\_Medio\% = 7,7\%$ e $3,8\%$ caso contrário). Existindo uma notória tendência ascendente quanto maior for a audiência média para o total da TV mais beneficiam os programas.

Para os períodos de menores audiências definidos no primeiro nível da Árvore (manhã, tarde e a partir da noite) os ratings são influenciados por $Tvr\_Prog\%$ (se $Tvr\_Prog\% > 14,65\%$, $Rat\_Prog\_Medio\% = 4,7\%$ e 2% , caso contrário). Para menores valores de $Tvr\_Prog\%$ ($< 14,65\%$) a variável $Tempo\_End$ destaca-se como favorecendo os programas com horas de conclusões superiores a 0,50, ou seja, às 12H da manhã.

Portanto, para este canal as audiências estão condicionadas essencialmente por factores de disponibilidade e pela capacidade deste canal pontualmente maximizar os seus *shares* face aos canais privados. Para períodos de menores audiências totais é relevante a hora de conclusão do programa. Os anteriores resultados correspondem aos dados associados ao primeiro nível programático. Apresentam-se alguns resultados considerando, dados parciais de ratings, ou seja, referentes a durações programáticas parciais e também por aplicação do Algoritmo CART.

➤ **Resultados atendendo à globalidade dos dados parciais, isto é, para $Rat\_Prog\%$ (P) (Árvore CART – software SAS)**

Do ponto de vista do planeamento comercial é necessário o desenvolvimento de modelos explicativos e de previsão de valores de ratings globais, mas também parciais, ou seja, neste último contexto considerando cada uma das partes que integram um programa. A variável critério corresponde a $Rat\_Prog\%$ (P) e as explicativas aos índices programáticos parciais e aos factores de planeamento. A Árvore apresenta uma capacidade explicativa de 83%. Recorrendo ao algoritmo CART, se incluirmos a identificação dos programas que constam da base de dados diferenciados em função da parte a que correspondem, para além das variáveis de caracterização dos programas e dos índices programáticos parciais, logo no primeiro nível da Árvore a identificação dos programas é a variável utilizada no seu desenvolvimento. Consequentemente, existe um número restrito de programas que se diferenciam dos restantes, apresentando maiores valores de índices programáticos.

Para o período em análise os maiores valores de ratings estão associados à principal telenovela da noite, à segunda parte do Jornal da noite da semana e aos programas directamente sucessores, inclusive ao fim-de-semana. Os programas de humor também conseguem maximizar as suas exposições. Correspondem ao períodos de prime time e do início da noite, sendo os programas de *lead_out* do espaço informativo favorecidos. No segundo conjunto de programas encontra-se o programa que antecede o espaço informativo da noite e a primeira parte do espaço

informativo, assim como a totalidade do espaço informativo da noite ao fim de semana, o espaço informativo da tarde de fim de semana, a segunda telenovela da noite em dias laborais e algumas partes de alguns filmes. Portanto, todos estes programas correspondem ao pré-prime-time, prime time e à noite.

Excluindo estes programas a média dos restantes programas é de 3,64% e estes são beneficiados quanto maior for o nível de *Tvr_Prog%*. Logo, os restantes programas diferenciam-se com base nos níveis das taxas médias de audiências para o total da TV.

Se excluirmos das análises a identificação dos programas são os níveis de *Tvr_Prog%* que condicionam os valores dos ratings e na ausência desta variável as variáveis *Tempo_Time* e *Tempo_End*. Os programas mais favorecidos com início antes das 19.30H correspondem aos horários compreendidos entre as 12H e as 15.40H, seguindo-se os programas da manhã. Uma vez mais os resultados estão condicionados por factores temporais. Se incluirmos na base de dados as *tipologias programáticas* para as folhas com maiores médias de ratings associadas uma vez mais surgem os factores temporais como sendo determinantes. Para *Tempo_End*<19.30H é que surgem as *tipologias*, sendo os ratings maximizados por espaços informativos e alguns filmes.

Se incluirmos em termos absolutos e relativos os índices programáticos parciais de caracterização da programação adjacente assim como os índices globais de caracterização da programação e aplicando o algoritmo CART

apenas com *Tvr_Prog%* e *Rat_Prog (T)%* consegue-se uma capacidade explicativa de 90%.

Para níveis de *Tvr_Prog%* inferiores aos 22,95% (*Rat_Prog%_Médio*=3,45% e 11,02% caso contrário) o que condiciona os valores dos ratings são os níveis de *Tvr_Prog%* e para os restantes obviamente que maiores valores globais estão também associados a maiores valores de *ratings* parciais.

Interessa sobretudo ressaltar que apenas para níveis mais elevados de *Tvr_Prog%* existe uma maior associação entre as audiências globais e parciais, ou seja, nos períodos de menores audiências existe uma maior disparidade de índices no decorrer de um programa. Nos períodos em que a audiência média para o total da TV é maior poderá existir uma maior estabilidade dos *ratings* no seu decorrer.

Excluindo das análises os *ratings* globais, (variância explicada de 93%) os programas são segmentados em função dos níveis de *Tvr_Prog%*.

Para *Tvr_Prog%*<7% os ratings estão mais associados aos *ratings* posteriores beneficiando da sua maior magnitude. Para valores de *Tvr_Prog%* no intervalo [7%, 9,7%] a influência advém dos níveis de *Tvr_Prog%*. Para valores no intervalo [9,75%, 14,25%] sobressai a influência de *Shr_Ldout%* e para valores no intervalo [14,25%, 33,35%] sobressai o indicador *Shr_Ldin%*.

Portanto, quanto maior a audiência disponível maior é a associação existente com os valores dos *shares* anteriores.

Para os níveis mais elevados de *Tvr_Prog%* (superiores aos 33,35%) em primeiro lugar os *ratings* associam-se a *Rat_Ldin%* e no caso de estes valores serem superiores a 9,35% será a *tipologia* do programa que condiciona os valores de *ratings*. As telenovelas, os *sketches* de humor, as séries de comédia são as tipologias que se destacam em maior grau face aos noticiários, concursos de beleza, passatempos e programas de apoio à comunidade.

Destes resultados verificamos que os níveis de *Tvr_Prog%* alteram a relevância da repercussão dos valores de *ratings* e dos *shares*, assim como dos valores dos programas adjacentes. Para os valores extremos de *Tvr_Prog%* é superior a associação entre os valores de *Rat_Prog%* (*T*) e os *ratings*; para os valores mais reduzidos sobressaem os *ratings* posteriores e para os valores mais elevados destacam-se os valores anteriores. Isto poderá corresponder ao início dos períodos com tendências ascendentes de valores de *ratings* e aos períodos de maiores audiências onde poderá existir um patamar de maior estabilidade das audiências e, consequentemente observamos uma maior associação com os valores dos *ratings* dos programas anteriores e, não, com os valores dos programas posteriores que vão na realidade decrescendo ao longo do tempo. Para níveis de *Tvr_Prog%* entre os 9,75% e os 33,35% os *ratings* são em maior grau função dos valores dos *shares* e uma vez mais para maiores níveis de *Tvr_Prog%* privilegiam-se os valores antecessores. Para níveis intermédios de *Tvr_Prog%* e que correspondem ao decorrer de períodos de transição existe uma maior dispersão de valores no seu decorrer, assumindo uma maior relevância os valores de *shares*. Nas tendências ascendentes de *ratings*

são mais relevantes os *ratings* posteriores e nas descendentes os *ratings* anteriores. Verificamos, também, que para os níveis mais elevados de *Tvr_Prog%* (<33,34%) e de *Rat_Ldin%* (>9,35%) sobressaem as tipologias programáticas, sendo favorecidas as telenovelas e os programas de humor (*Rat_Prog%_Médio*=14,7%) e em menor grau os espaços informativos e as restantes tipologias enumeradas. Correspondem a programas que estão intercalados entre o espaço informativo e a telenovela da noite ou inclusive no decorrer da telenovela. Estes programas saem beneficiados por estarem inseridos em períodos de maiores audiências, sendo tanto maiores os valores dos seus *ratings* quanto maior forem os valores antecessores e pela proximidade face à primeira telenovela da noite. A maioria das vezes os seus *ratings* são inferiores aos da telenovela, apesar da sua proximidade.

Excluindo, os valores dos *ratings* e de *shares* adjacentes os *ratings* estão condicionados por *Tvr_Prog%* e para os maiores valores e superiores a 33,35% por parte do horário de conclusão, o que denota a penalização que sofrem os programas no decorrer da noite por finalizarem em horários cada vez mais tardios. Rapidamente assistimos a um gradual abandono das audiências. Desenvolvendo um pouco mais a Árvore para níveis de *Tvr_Prog%* inferiores aos 7%, os *ratings* são superiores ao fim-de-semana onde se transmite programas juvenis, em vez de outros programas matinais. Para o espaço informativo da noite os *ratings* são favorecidos por maiores níveis de *Tvr_Prog%*, sendo este factor responsável pelos maiores valores associados à segunda parte do espaço informativo.

Posteriormente e até à conclusão da principal telenovela da noite destaca-se este programa face aos restantes e a partir daí os *ratings* estão influenciados pelos níveis de *Tvr_Prog%*.

➤ **Resultados considerando a globalidade dos dados parciais desde o espaço informativo da noite e tendo como variável dependente *Rat_Prog (P)* (CART - SAS)**

Restringindo as análises aos períodos que decorrem desde o espaço informativo da noite, uma vez mais será *Tvr_Prog%* que diferencia os programas no primeiro nível da Árvore e no segundo nível os valores de *ratings* globais.

Quando os *ratings* globais são inferiores a 3,55% e *Tvr_Prog%* < 23%, portanto no final da noite e após as 23.30H, os *ratings* parciais são superiores quando também o são para os programas/partes anteriores e posteriormente as partes posteriores.

Para *ratings* globais superiores aos 3,55% e iguais níveis de *Tvr_Prog%*, portanto abrangendo parte da segunda telenovela e a terceira telenovela da noite, os *ratings* parciais são tanto maiores quanto maiores forem os *ratings* globais. Para níveis de *Tvr_Prog%* mais elevados (>23.33%) quanto maiores forem os *ratings* globais (>11.6%) mais estes valores estão associados aos valores parciais e para os restantes valores de *Rat_Prog% (T)* associam-se com os valores de *shares*. Os programas muito populares apresentam uma maior semelhança entre as audiências parciais e as globais e os programas que apresentam uma menor exposição conseguem maximizar os seus *ratings* para as partes que também maximizam os valores de *shares*.

Neste último contexto os *ratings* são tanto maiores quanto maiores forem os níveis de *Tvr_Prog%*. Se excluirmos os *ratings*, as taxas médias de audiências para o total da TV, os *shares* dos programas e os seus valores adjacentes as variáveis que assumem uma maior relevância são as horas de conclusão dos programas e, posteriormente as horas de início. Para as horas de conclusão mais tardias é relevante a fidelidade aos programas posteriores que quanto mais tarde se prolongarem a nível global mais os beneficiam.

Excluindo os valores de fidelidade o que condiciona em primeiro lugar os *ratings* são as horas de conclusão dos programas. Quando a conclusão do programa for anterior às 24H, as horas de início associadas ao pré-prime-time e ao início do prime time penalizam os *ratings* e neste contexto os programas diferenciam-se com base em tipologias e dias da semana, sendo superiores ao Sábado, Sexta e Segunda.

Para as horas de início que correspondem ao decorrer da segunda telenovela da noite, os maiores valores de *ratings* estão associados às partes programáticas que não correspondem às primeiras partes de qualquer programas. Para os restantes programas quanto maior for a duração parcial face à duração global maiores são os valores de *ratings*. Entre os espaços informativos e a segunda telenovela da noite os programas aos quais corresponde uma única parte programática maximizam os valores dos *ratings*.

➤ **Resultados considerando a globalidade dos dados parciais a partir de pré-prime-time e tendo como variável dependente *Rat_Prog%* (P) (CART - SAS)**

Considerando os dados a partir de pré-prime-time se incluirmos os nomes dos programas e a programação concorrente, mantém-se as variáveis *Tvr_Prog%* e *Rat_Prog%* (T), como sendo as mais relevantes para a previsão dos valores de *ratings* parciais. Quanto mais elevadas forem as taxas médias de audiências para o total da TV maior é a associação existente entre os valores dos *ratings* parciais e os valores globais enquanto que para os valores mais reduzidos de *Tvr_Prog%* será esta a variável que mais os condiciona, assim como a hora de início da programação.

➤ **Resultados considerando a globalidade dos dados parciais a partir de pré-prime-time e tendo como variável dependente *Tvr_Prog%* (P) (CART - SAS)**

Se com os dados parciais quisermos avaliar como variável dependente *Tvr_Prog%* e utilizando como dados a totalidade da emissão diária, a partir do programa de pré-prime-time até às 24H, os seus valores são função da hora de conclusão do programa e para os restantes em primeiro lugar surgem as tipologias.

Nos períodos de maiores audiências, os factores temporais penalizam as audiências e é neste período que estes valores podem ser maximizados num horizonte temporal mais alargado. Para os restantes períodos, uma vez que existe uma selecção mais pontual os valores de *Tvr_Prog%* diferenciam-se com base nas tipologias a que estão associadas.

Os valores são maximizados por parte dos espaços informativos, alguns filmes, galas, acontecimentos especiais e algum concurso. O segundo nó da Árvore com maiores valores de *Tvr_Prog%* também inclui a maioria das tipologias associadas a filmes, séries e telenovelas.

Para as restantes tipologias são as horas de início. Quanto mais tarde for a hora de conclusão mais beneficiados são os programas. Portanto, ao longo da maior parte da emissão diária se excluirmos os programas pontualmente mais atractivos os restantes são beneficiados ao nível das audiências totais desde que estejam associados a horas de início mais tardias. Se excluirmos as horas e tempos de início de conclusão das análises os valores de *Tvr_Prog%* são diferenciados em função das tipologias e as que maximizam as audiências totais (*Tvr_Prog%_Médio*=23%) têm valores tanto maiores quanto menores forem as saturações publicitárias globais (*Sat_Prog%*<17%) (Média de *Tvr_Prog%* no valor de 28% e 18% caso contrário). Para as partes que estão inseridas em programas com maiores saturações beneficiam em termos de *Tvr_Prog%* as partes que são antecedidas por intervalos de menor duração (Média de *Tvr_Prog%*=24% e 16,4% caso contrário). Para as tipologias a que estão associados menores valores de *ratings* (*Rat_Prog%*<9,2%) um subconjunto de programas diferencia-se com base nos dias da semana e inclui séries e filmes essencialmente, sendo estes superiores ao fim-de-semana, Sexta e Segunda (Média de *Tvr_Prog%*=15% e 9,8% caso contrário).

O outro subconjunto de programas com valores mais reduzidos de *Tvr_Prog%* (7%) diferencia-se com base na duração dos programas sucessores, beneficiando com os maiores valores (*Média de Tvr_Prog%*=12,5% e 6% caso contrário). Se restringirmos as análises aos programas com início após o pré-prime-time os valores de *Tvr_Prog%* são função das horas de início e em maior grau de conclusão dos programas.

➤ **Resultados atendendo à globalidade dos dados programáticos parciais do canal estatal tendo como variável dependente *Rat_Prog%* (P) (CART - SAS)**

Repetindo as análises programáticas parciais para o canal estatal RTP1 e incluindo a totalidade das variáveis (índices programáticos globais e parciais e valores relativos, índices parciais dos programas adjacentes em termos absolutos e relativos, variáveis de identificação da programação e sua caracterização, variáveis temporais a nível global e parcial e sua inter-relação) a variável utilizada no primeiro nível da Árvore é o *nome do programa* com a identificação da parte que lhe corresponde. Esta Árvore apresenta 87% de variância explicada.

Os maiores valores de *Rat_Prog%* correspondem aos seguintes programas: aos espaços informativos da noite em dias laborais e apenas à primeira parte do espaço informativo da tarde de dias laborais, o concurso “Preço Certo” que antecede o espaço informativo da noite, o concurso de prime time, alguns jogos de futebol (Taça UEFA, Super Liga), programas de Fórmula I, eventos como o 1º Maio, Festival da Canção, Procissão das Velas e Missa benção dos doentes.

Estes programas, por sua vez, diferenciam-se em função dos valores de *shares* (*Shr_Prog%*>54%, *Rat_Prog%_Médio*=24,6% e 9,3% caso contrário). Para os menores valores de *shares* existe associação entre os valores globais e os valores parciais de *ratings*.

Para os restantes programas, estes por sua vez subdividem-se em dois grupos com as maiores médias associadas aos programas de *lead in* e de *lead out* aos espaços informativos da tarde, jogos de futebol, a segunda parte do Jornal da Tarde, alguns filmes, galas, Fórmula 1 e especiais desporto. Estes programas diferenciam-se em função dos níveis de *Tvr_Prog%* a que correspondem (se *Tvr_Prog%*>35%, *Rat_Prog%_Médio*=22% e 5% caso contrário).

Os restantes programas diferenciam-se, uma vez mais em função dos programas a que correspondem. Os programas mais penalizados são os matinais, os juvenis, os programas da tarde e os filmes (*Rat_Prog%_Médio*=1,5%).

Estes resultados permitem confirmar as tendências apontadas anteriormente reforçando a ideia de uma selecção mais pontual neste canal estatal. Consequentemente, a identificação dos programas corresponde à variável mais relevante para a segmentação da amostra, em vez dos valores de *Tvr_Prog%* ou de *Ratings*.

Os programas que maximizam as audiências diferenciam-se em função da sua capacidade para maximizar os seus *shares*. Os programas mais penalizados a este nível ($Shr\_Prog\% < 54\%$) apresentam *ratings* parciais tanto superiores quanto maiores, também, forem os *ratings* globais. Portanto, para os programas de maior popularidade existe uma certa estabilidade no seu decorrer aproximando os valores parciais dos valores globais.

O segundo subconjunto de programas com maiores índices de audiências diferenciam-se em função dos níveis de $Tvr\_Prog\%$, uma vez que englobam programas repartidos ao longo de toda a emissão diária. Verificamos que os espaços informativos correspondem a picos de audiências favorecendo unicamente os programas adjacentes e, posteriormente os *ratings* entram em declínio. Os restantes programas diferenciam-se em função do programa a que correspondem.

Se excluirmos o nome dos programas para este canal a variável utilizada no primeiro nível da Árvore é $Rat\_Prog\% (T)$, em vez das variáveis $Tvr\_Prog\%$ ou $Rat\_Ldin\%$ que assume um maior destaque para os canais privados. Para este canal existe uma maior estabilidade nas audiências no decorrer de um programa. Para $Rat\_Prog\% (T)$ superior aos 6%, os programas diferenciam-se em função dos valores dos *shares* e para os menores valores ($Shr\_Prog\% < 54\%$) em função de $Tvr\_Ldout\%$ (se $Tvr\_Ldout > 26\%$ $Rat\_Prog\%_{Médio} = 10\%$ e 6,2% caso contrário).

Portanto, para os programas com maiores índices de audiências, os *ratings* são tanto superiores quanto maiores forem os valores dos *shares* (válido essencialmente para programas desportivos e para outros que pontualmente também sejam muito atractivos). Para *shares* de audiência inferiores aos 54% os programas saem beneficiados por uma maior proximidade face a programas posteriores aos quais correspondam maiores índices de audiências totais. Portanto, para este canal o efeito *lead out* é mais relevante do que o efeito *lead in*, ou seja, existe uma maior associação entre os programas e os valores dos programas posteriores do que face aos programas anteriores.

Para valores de $Rat\_Prog\% (T)$ inferior aos 6% os *ratings* são tanto maiores quanto maiores forem os valores de $Tvr\_Prog\%$. Para valores de $Tvr\_Prog\% > 10\%$, os programas diferenciam-se com base nos valores de $Rat\_Ldout\%$. Uma vez mais também para os programas de menores audiências é mais relevante o efeito *lead out*. Os espaços informativos beneficiam os programas anteriores e funcionam como *lead out* para estes. O concurso da noite beneficia os programas que são transmitidos entre ele e o espaço informativo funcionando uma vez mais como *lead out*. Após os programas mais populares verifica-se uma rápida redução nos índices de audiências não existindo, portanto as tendências de *lead in*.

Excluindo as variáveis enumeradas nos dois parágrafos anteriores com excepção de $Rat\_Prog\% (T)$ esta é a variável mais utilizada no desenvolvimento da Árvore.

Para valores de *Rat_Prog%* (*T*) inferiores aos 6% existe associação entre os *ratings* parciais e valores de *Tvr_Prog%*. Neste contexto os *ratings* são função dos períodos horários que cada uma das partes programáticas abrangem, sendo obviamente favorecidas as que se aproximam dos espaços informativos. Para valores de *Tvr_Prog%* superiores aos 10% as partes de um programa que saem beneficiadas apresentam audiências 3,42 vezes superiores aos valores dos *ratings* globais.

Nestes casos existe uma forte desigualdade nos valores dos *ratings* no decorrer de um programa. Para valores de *Rat_Prog%*>6%, uma vez mais os programas subdividem-se em função dos *ratings* globais e posteriormente em função do quociente entre *Rat_Prog%* (*P*) e *Rat_Prog%* (*T*). Os valores dos *splits* são de 0,90 para este quociente quando *Rat_Prog%* (*T*)>14% e de 0,83 para *Rat_Prog%* (*T*)<14%. Portanto, para as maiores audiências globais os *ratings* parciais são proporcionalmente mais elevados.

Excluindo a variável *Rat_Prog%* (*T*) assume uma maior importância a hora de conclusão do programa e os maiores valores verificam-se para *Tempo_End_Out*>0,84. Para estes programas e com uma hora de conclusão a partir das 20H, os programas diferenciam-se em função de *Atv*, ou seja, da permanência média das audiências no seu decorrer. Para os restantes horários de conclusão existe uma maior associação entre os *ratings* e as tipologias, sendo os *ratings* maximizados para os programas desportivos, os noticiários, os concursos e as informações de acontecimentos especiais.

Para estas tipologias os programas são beneficiados por maiores valores de *Rat_Ldin%* e para as restantes tipologias por horários de início mais tardios.

➤ **Resultados atendendo aos dados programáticos parciais do canal estatal RTP1 e considerando os dados a partir das 19.30H (CART - SAS)**

Considerando apenas os dados a partir das 19.30H verificamos que os programas são em primeiro lugar diferenciados pelo respectivo nome. Os *ratings* minimizam-se para inúmeros filmes, programas de debate, documentários sendo o valor dos seus *ratings* função dos horários de início dos programas (Para *Hora_Início*>22.30H corresponde um *Rat_Prog%_Médio*=3,6% e 6,7% caso contrário).

Para os programas que maximizam os *ratings* (por exemplo, telejornal, jogos de futebol, concurso de prime-time, “o preço certo”, festival da canção, ...) os programas diferenciam-se em função dos seus *shares* de audiências. Para valores de *shares*>54% corresponde um *Rat_Prog%_Médio*=25,6% e para os restantes a variável mais associada à variável em estudo são os valores dos *ratings* globais (*Rat_Prog%* (*T*)>10% corresponde um *Rat_Prog%_Médio*=12% e 9% caso contrário). Portanto, os programas com menores níveis de exposições diferenciam-se em função dos horários e os restantes programas em função dos valores dos *shares* e das audiências globais. Para os programas mais populares é maior a associação entre os *ratings* parciais e os globais do que com os valores dos *ratings* dos programas adjacentes.

Não existe proximidade de valores entre os programas próximos, podendo existir uma considerável perda ou acréscimo na dimensão das audiências de programa para programa, mas com uma certa estabilidade no seu decorrer. Os programas da noite com início a partir das 22.30H vão perdendo audiências de forma acentuada no decorrer da noite, sendo, portanto as audiências minimizadas pelos programas transmitidos em horários mais tardios.

Se excluirmos a identificação dos programas prevalece a influência dos valores de *shares* e dos *ratings* globais. Portanto, uma vez mais os programas diferenciam-se em função da sua capacidade de maximização dos seus *shares* e, não, em função dos factores temporais.

Se eliminarmos os programas que geram observações atípicas e incluirmos a totalidade das variáveis, os programas diferenciam-se em função dos valores de *ratings* globais e para os menores valores e inferiores a 10%, também em função dos valores de *shares*. Portanto, excluindo os programas desportivos que geram grandes audiências quanto maiores forem as audiências globais maiores são também as audiências parciais e maior será a associação entre estas duas variáveis. Para os restantes programas os *ratings* são tanto maiores quanto também o forem os valores de *shares*, uma vez que existe uma maior separação entre os índices globais e os índices parciais.

Excluindo os *ratings* globais e os valores de *shares*, destacam-se algumas tipologias de programas, como, por exemplo, os noticiários, os concursos e os jogos de futebol que maximizam os valores dos *ratings*.

Nestes casos existe associação com *Rat_Ldin%* e para menores valores desta variável com *Tvr_Ldout%*. Os programas que maximizam as audiências ou que também sucedem programas com valores de *ratings* superiores aos 14% (*Rat_Prog%_Médio*=16,4%) ou que sucedem programas com *ratings* inferiores aos 14%, mas com *Tvr_Ldout%*>35,5% (*Rat_Prog%_Médio*=12%).

Portanto, as audiências estão condicionadas por factores temporais e por proximidade face aos programas com maiores níveis de exposições. Para as restantes tipologias os programas beneficiam por uma maior proximidade face a programas que apresentam maiores valores de fidelidade e superiores aos 44% (8,7% de *Rat_Prog%_Médio* e 6,38% caso contrário). Para valores de *Tvr_Prog%*>40%, se a hora de início for anterior às 21H os *ratings* são superiores (25% de *Rat_Prog%_Médio* e 11% caso contrário). Para *Tvr_Prog%* inferior aos 40% será a hora de conclusão de um programa que condiciona os valores dos *ratings*, sendo favorecidos por horários anteriores às 22.30H e neste caso essencialmente os programas que apresentam apenas uma parte. Para horários de conclusão após as 22.30H será a sua duração que condiciona os *ratings*.

Resultados atendendo à globalidade dos dados programáticos parciais do canal TVI (CART - SAS)

Por último para o canal TVI e incluindo a totalidade das observações e das variáveis de audimetria, os valores dos *ratings* parciais estão mais associados a valores de

Tvr_Prog_Ldin% (a segunda melhor variável seria *Tvr_Prog%*). Quanto maior a audiência média para o total da TV maiores também são os valores dos *ratings*.

Para valores de *Tvr%* superiores aos 20% existe uma maior associação entre os *ratings* parciais e os *ratings* globais. Para valores de *ratings* globais superiores aos 11% os *ratings* são superiores para os programas que sucedem programas com uma maior duração (>9700s). Para valores de *ratings* globais inferiores aos 11% os programas beneficiam com maiores valores de *Rat_Ldin%* (>8%).

Para valores de *Tvr_Ldin%*<20%, os programas diferenciam-se em função de *Tvr_Prog%* que quando é superior a 12% permite uma maior associação entre os *ratings* parciais e os valores globais.

Portanto, também para a TVI o que condiciona os *ratings* corresponde em primeiro lugar aos factores de disponibilidade da audiência que quanto maior permite uma maior aproximação entre os *ratings* globais e os valores parciais.

Excluindo estas variáveis no primeiro nível da Árvore surge *Rat_Ldin%* (se *Rat_Ldin%*>8,8% *Rat_Prog%_Médio*=11% e 3,2% caso contrário). Os programas beneficiam com uma maior proximidade face aos programas que apresentam uma maior exposição. Neste caso a *hora de conclusão* condiciona os *ratings* beneficiando aqueles que antecedem as 20.45H. Para valores de *Rat_Ldin%*<8,8% será a hora de início que os beneficia desde que seja após as 18.30H e neste caso por uma maior proximidade a maiores valores de *Rat_Ldin%*.

Para horários de início anteriores às 18.30H, a variável *Rat_Idout%* diferencia os programas. Fora dos períodos de maiores audiências, os programas beneficiam por sucederem a outros com maior níveis de exposição (para horas de início posteriores às 18.30H) ou anteriores (para horas de início anteriores às 18.30H).

Portanto, para os maiores valores de *Rat_Ldin%* são os factores temporais que condicionam os *ratings*, favorecendo os programas até as 20.45H. Para menores valores de *Rat_Ldin%* o que corresponde aos programas no decorrer da noite, após o espaço informativo da tarde e aos períodos de tendências ascendentes será a hora de início de um programa que condiciona os *ratings*. Após as 18.30H quanto mais populares forem os programas anteriores mais são beneficiados os programas.

Por exemplo, no decorrer da noite em parte devido aos efeitos de herança os programas que sucedem outros programas mais populares saem beneficiados e nos restantes períodos em que *Rat_Ldin%* é inferior aos 8,8% por uma maior proximidade face ao início de programas com maiores níveis de exposições. Ao longo de todo o dia até ao final da tarde, ou seja, até às 18.30H existe uma selecção mais pontual, por conseguinte apenas os programas que antecedem os programas com uma maior notoriedade é que apresentam maiores valores de *ratings*. Uma vez concluídos os programas de maior interesse não existe duplicação de audiências para períodos posteriores e consequentemente não existe associação com *Rat_Ldin%*.

Ignorando as variáveis relativas aos *ratings*, valores de *Tvr_Prog%* e valores de *shares*, os programas diferenciam-se em função do horário de início, sendo favorecidos os que começam após as 20.45H e em especial antes das 22.45H.

Portanto, confirma-se que para os programas de prime time e noite os seus *ratings* são função de factores temporais. Verificamos que os grandes intervalos de variabilidade das telenovelas da noite eram devido às oscilações existentes nos horários e que têm uma grande repercussão na dimensão das audiências.

Para os restantes horários sobressaem os espaços informativos da tarde (FS), o espaço informativo da noite, ao longo de toda a semana, as telenovelas antes do espaço informativo da noite, as três sessões de cinema ao Domingo, a Taça de Portugal e algumas matinés de cinema ao Sábado.

Obviamente que destes programas destacam-se os jogos de futebol (*Rat_Prog%_Médio*=21% e 7,5% caso contrário). Para os restantes programas será o nível de *Tvr_Prog%* (*T*) que condiciona os *ratings* (Se *Tvr_Prog%*>11%, *Rat_Prog%_Médio*=3,5% e 2% caso contrário).

Portanto, para a TVI nos períodos de maiores audiências é fundamental o horário de início. Para o resto do dia existe um subconjunto de programas que se destaca, sobressaindo os eventos desportivos e para os restantes programas o valor de *Tvr_Prog% global* determina os seus níveis de audiências.

➤ **Dados referentes a *shares* programáticos do Canal SIC e considerando a totalidade da emissão diária e períodos mais restritos (CART - SAS)**

Por último, resta-nos avaliar os valores de *shares* de audiências.

Se nos restringirmos aos programas a partir do pré-prime-time os valores são segmentados em função de horas de início e de conclusão dos programas. Em alguns dos anos, os *shares* são maximizados entre a segunda parte do espaço informativo até à conclusão da primeira telenovela da noite e posteriormente segue-se o início do espaço informativo.

Para os restantes períodos horários, os *shares* são tanto maiores quanto mais breves forem as horas de conclusão sendo maximizados para os que programas que finalizam antes das 24H.

Considerando, a totalidade da emissão a variável que permite segmentar os programas em função dos valores de *shares* são os níveis de *Tvr_Prog%* ou do *nome do programa*, as variáveis que se utilizam no primeiro nível da Árvore. Os programas com maiores valores de *shares* (*Shr_Prog%_Médio*=40%) diferenciam-se em dois subconjuntos, sendo os *shares* tanto maiores quanto maior for a fidelidade que lhes é dirigida.

Os *shares* são máximos (*Shr_Prog%_Médio*=48%) para os espaços matinais de fim-de-semana de programas juvenis e para algumas partes dos espaços destinados à transmissão de filmes.

Seguem-se algumas partes de espaços destinados à transmissão de filmes, as últimas partes do programa matinal, a segunda parte e o programa de *lead_out* do espaço informativo da tarde, o Primeiro Jornal de fim-de-semana, a telenovela ou programas de *lead_out* do espaço informativo da noite (*Shr_Prog% Médio*=37%). Para os programas com menores valores de *shares*, uma maior proximidade face a maiores valores de *Fid_Ldout* contribui para uma maximização dos valores de *shares*.

Os *shares* estão associados em maior grau com os valores de fidelidade. Eliminando as variáveis associadas à fidelidade da audiência assumem maior relevância os *ratings* e os níveis de *Tvr_Prog%*.

Destes resultados concluímos que as Árvores segmentam os programas por ordem de desempenho, utilizando determinadas variáveis explicativas e que quase sempre correspondem aos factores temporais. Contudo, não nos elucidam a respeito de quais são os factores que condicionam a variabilidade observada em programas em concreto, e em quais são os valores entre os quais oscilam.

➤ Síntese da capacidade de explicativa das análises de Regressão Não Linear

Nos seguinte *Tabela F.6.4.1* resume-se a capacidade explicativa dos modelos de Regressão não linear calculados. Identifica-se a metodologia, amostra e as variáveis incluídas em cada modelo.

Algoritmo	Amostra	% Variância Exp.	Variáveis
CHAID	Totalidade das Observações	87%	Tvr_Prog% Rat_Ldin% Rat_Ldout% Tipologias
CART	Totalidade das Observações	82%	Tvr_Prog% Rat_Ldin% Rat_Ldout% Tipologias
CART	Totalidade das Observações (Excluindo atípicas)	80%	Tvr_Prog% Rat_Ldin% Período
CART	Totalidade das Observações	84%	Rat_Ldin% Período Tipologias Tvr_Prog%

Algoritmo	Amostra	% Variância Exp.	Variáveis
CART	Canais Privados pré prime time, prime time e noite	80%	Hora Conclusão do Programa Tvr_Prog% Rat_Ldin% Fid_Prog% Canal Tipologias Rat_Ldout%
CART	Canais Privados	90%	Tvr_Prog% Período Horário Fid_Prog% Semana ou FS Canal Tipologias
CART	RTP1	80%	Tvr_Prog% Período Hora_Conclusão Shr_Prog%
CART	Canais Estatais	80%	Período Horário Tvr_Prog% Hora_Conclusão

Tabela F.6.4.1 – Capacidade explicativa das Árvores de Regressão

V. Análises de Regressão não linear Recorrendo a Redes Neurais

Nas seguintes análises utilizaram-se as *Redes Neurais MLP* unidireccionais de aprendizagem supervisionada e as *Redes Neurais RBF* unidireccionais de tipo híbrido. Os aspectos a definir nestas Redes Neurais correspondem ao número de centros, seu posicionamento, a função de activação, os parâmetros da função e os pesos associados à capa final. Considerando as *Redes Neurais MLP* e considerando três das variáveis explicativas mais relevantes segundo a Regressão Linear Multivariada Simbólica e as Árvores de Regressão, ou seja, *Tvr_Prog%*, *Rat_Ldin%* e *Rat_Ldout%* a melhor Rede MLP corresponde a uma Rede Neuronal de tipologia 3-2-1 na qual as variáveis de entrada se estandardizaram e utilizou-se o algoritmo Gradiente Conjugado.

Os factores de normalização são para *Tvr_Prog%* ($\mu=20.56$, $\sigma=11.32$), *Rat_Ldin%* ($\mu=6.33$, $\sigma=4.25$), *Rat_Ldout%* ($\mu=6.27$, $\sigma=4.41$) e *Rat_Prog%* ($\mu=6.4$, $\sigma=4.22$). A função de activação para a capa oculta é a tangente hiperbólica e para a capa de saída é a função linear.

Da Rede MLP constam um total de onze pesos a calcular e que apresentam os valores que constam do *Tabela F.6.4.2*, *Tabela F.6.4.3* e do *Figura F.6.4.2*. Os pesos iniciais têm valores no intervalo $[-0.8, 0.8]$ segundo a distribuição Uniforme. O problema associado às variáveis *Rat_Ldin%* e *Rat_Ldout%* consiste em que *a priori* também não se conhecem os seus valores, e, logo, do ponto de vista da previsão não são muito interessantes.

Neuronas Associadas	De	Para	Pesos	Sinal
Rat_Ldin -> H11	Rat_Ldin	H11	0,530	+
Rat_Ldout -> H11	Rat_Ldout	H11	0,045	+
Tvr -> H11	Tvr	H11	0,100	+
Rat_Ldin -> H12	Rat_Ldin	H12	- 0,635	-
Rat_Ldout -> H12	Rat_Ldout	H12	0,500	+
Tvr -> H12	Tvr	H12	1,040	+
BIAS -> H11	BIAS	H11	- 0,250	-
BIAS -> H12	BIAS	H12	0,056	+
H11 -> Rat	H11	Rat	5,550	+
H12 -> Rat	H12	Rat	2,510	+
BIAS -> Rat	BIAS	Rat	7,290	+

Tabela F.6.4.2 – Pesos da Rede Neuronal MLP

Indicadores	Treino	Validação	Teste
Average Squared Error	4,7	5,068	4,3
Root Average Squared Error	2,17	2,2	2,1
Root Final Prediction Error	2,2		
Root Mean Squared Error	2,2	2,2	2,1
Mean Squared Error	4,8	5,068	4,3
Maximum Absolute Error	5,9	1,3	5,98
Final Prediction Error	4,9		

Tabela F.6.4.3 – Grau de Ajustamento da Rede Neuronal MLP

A Rede MLP apresenta um MSE de 4,797 para a amostra de treino e de 5,068 para a amostra de validação. Discretizando os *ratings* programáticos em intervalos de 4% de amplitude verificamos que 82.13% de observações foram correctamente classificadas. Portanto, para estas observações os seus valores estimados não diferem dos reais em mais de 4%.

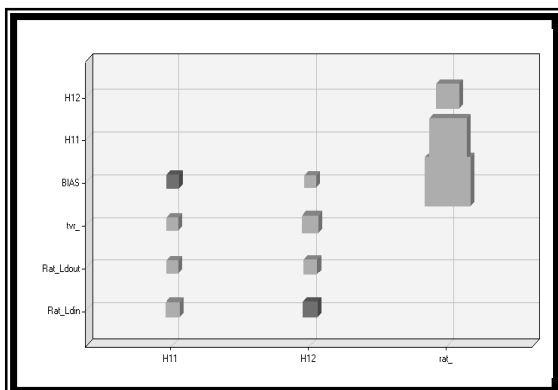


Fig. F.6.4.2 – Pesos da Rede Neuronal MLP

Utilizando diferentes combinações das restantes variáveis explicativas, mas conhecidas no momento de realizar a previsão o melhor modelo é o que inclui as variáveis *Tvr_Prog%*, *Bsh_Prog%* e *Sat_Prog* (Fig. F.6.4.3, Tabela F.6.4.4 e Tabela F.6.4.5).

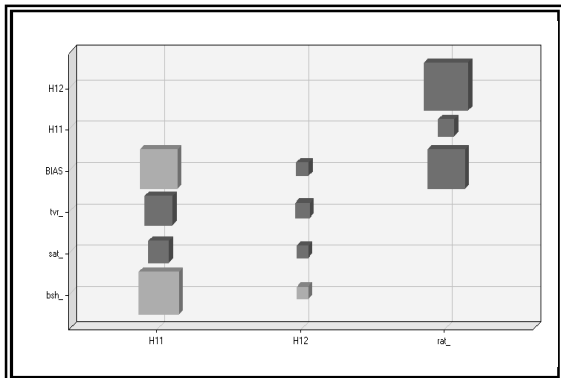


Fig. F.6.4.3 – Pesos da Rede Neuronal MLP

Esta Rede MLP (3-2-1) apresenta um MSE de 2.086 para a amostra de treino e de 3,094 para a amostra de teste. Considerou-se o algoritmo gradiente descendente e pesos iniciais uniformes no intervalo $[-0.8, +0.8]$. A função de activação é a tanh para a capa oculta.

Consequentemente, as Redes Neurais MLP apresentam resultados satisfatórios para o desenvolvimento de modelos de previsão de valores de *ratings*, ainda que não permitam “explicar” as interações existentes, sem recorrer a abordagens complementares.

De	Para	Pesos	Sinal
Bsh	H11	-6,280	-
Sat	H11	1,996	+
Tvr	H11	3,554	+
Bsh	H12	0,000	+
Sat	H12	0,009	+
Tvr	H12	0,595	+
BIAS	H11	-5,666	-
BIAS	H12	0,171	+
H11	Rat	0,965	+
H12	Rat	7,254	+
BIAS	Rat	5,660	+

Tabela F.6.4.4 – Pesos da Rede Neuronal MLP

Indicadores	Treino	Validação	Teste
Average Squared Error	2,040	3,090	2,099
Root Average Squared Error	1,428	1,759	1,448
Root Final Prediction Error	1,460	.	.
Root Mean Squared Error	1,444	1,759	1,440
Mean Squared Error	2,086	3,094	2,099
Maximum Absolute Error	6,645	13.17	6,335
Final Prediction Error	2,133	.	.

Tabela F.6.4.5 – Grau de Ajustamento da Rede Neuronal MLP

Considerando as Redes Neurais RBF, ou seja, com funções de base radial que são mais rápidas no seu treino do que as Redes MLP e experimentando diferentes combinações das variáveis explicativas a melhor Rede Neuronal RBF inclui as variáveis: *Tvr_Prog%*, *Bsh_Prog%*, *Sat_Ldin%* e considerando como variável dependente *Rat_Prog%*.

Experimentando-se diferentes tipologias de Rede e variando o número de centros (nós na capa oculta), a melhor Rede corresponde a três centros cujos centroides foram calculados utilizando o algoritmo de K Média.

Definidos os centroides calcula-se o parâmetro de escala da função de transferência (a Gaussiana) que corresponde à média das distâncias quadráticas entre os clusters, neste caso com o valor de 800. O parâmetro de normalização mede a largura da função Gaussiana e equivale ao raio da influência da neurona no espaço das entradas. Esta função aplica-se à distância euclidiana que separa o vector de entradas do seu centroide.

O valor dos centros é o seguinte:

- (1) Centro 0 - +0.073 (*Tvr*)/ +1.193 (*Bsh*) / -0.93 (*Sat_ldin*);
- (2) Centro 1 - +0.80/ -0.18/ 0.93;
- (3) Centro 2 - +0.30/ -0.21/ -0.93.

A Rede RBF apresenta um MSE de 1.99 para a amostra teste com 30% das observações e um erro absoluto médio no valor de 1.33.

Uma vez mais se discretizarmos os valores dos *ratings* programáticos em intervalos de 4% de variabilidade, 79.8% de observações são correctamente classificadas no mesmo intervalo de valores da variável de saída.

A amostra de treino que representa 70% das observações apresenta um MSE de 2.16 e um MAE de 1.41 e obtemos 78.6% de observações correctamente classificadas na mesma situação. O RMSE calcula-se através da fórmula representada por

$$RMSE = \sqrt{\frac{1}{n} \sum_{x \in D} (h(x) - f(x))^2}$$

e o MAE representado através da fórmula

$$MAE = \frac{1}{n} \sum_{x \in D} |h(x) - f(x)|$$

Uma vez mais os resultados do ponto de vista da previsão são razoáveis, atendendo aos valores que constam da bibliografia da especialidade internacional.

Na seguinte *Tabela F.6.4.6* consta uma síntese da capacidade predictiva das Redes Neurais implementadas neste capítulo.

Incluindo variáveis que são conhecidas no momento de realizar a previsão a percentagem de variância explicada para a amostra de teste aproxima-se dos valores alcançados pelas Árvores de Regressão e das abordagens lineares. A capacidade explicativa para a totalidade das abordagens aproxima-se dos 80%.

Algoritmo	Amostra	Capacidade Predictiva	Variáveis
MLP	Totalidade das Observações	82.13% em análises de classificação para amostra de teste MSE de 4,797 para a amostra de treino e MSE de 5,068 para a amostra de validação	Tvr_Prog% Rat_Ldin% Rat_Ldout%
MLP	Totalidade das Observações	MSE de 2.086 para a amostra de treino e MSE de 3,094 para a amostra de teste	Tvr_Prog% Bsh_Prog% Sat_Ldin%
Algoritmo	Amostra	Capacidade Predictiva	Variáveis
RBF	Totalidade das Observações	MSE de 1.99 para a amostra de teste e MSE de 2,16 para a amostra de treino 79,8% em análises de classificação para amostra de teste	Tvr_Prog, Bsh_Prog, Sat_Ldin

Tabela F.6.4.6 – Síntese de capacidade explicativa e de previsão das Redes Neurais MLP e RBF

F.7. A EXTRACÇÃO DE REGULARIDADES ESTATÍSTICAS: AS ANÁLISES EXPLICATIVAS E DE PREVISÃO DAS AUDIÊNCIAS COMERCIAIS

F.7.1. INTRODUÇÃO

No passado tem-se frequentemente questionado a publicidade televisiva quanto à sua eficácia. No entanto, a televisão mantém-se e será no futuro um dos *media* efectivamente mais utilizado para este fim (*Rubinson, 2009; Sharp, Beal & Collins, 2009*). As tecnologias de informação e as alternativas de entretenimento, como, por exemplo, a Internet não têm conseguido afastar os telespectadores do meio televisivo (*Cooper & Tang, 2009; Green, 2007*), aliado à sua reduzida penetração nos lares nacionais (*Fernandes, 2008*).

Identificamos inúmeras áreas de análise no campo da publicidade televisiva, sendo as temáticas abordadas muito diversificadas. Os trabalhos abrangem múltiplas fontes de dados e métodos de análise, existindo uma maior incidência em trabalhos de cariz qualitativo ou recorrendo aos inquéritos. O contínuo acréscimo dos custos publicitários aliado à fragmentação das audiências e à necessidade de desenvolver estratégias de Marketing direccionadas a determinados nichos de mercado, realçou o interesse dos profissionais do meio quanto ao desenvolvimento de modelos explicativos das audiências comerciais, assim como na quantificação dos futuros níveis de audiências dos suportes publicitários.

O centro de investigação mundial, WARC, na área dos *media* apresenta uma categorização das temáticas em análise, referindo-se no *Quadro F.7.1.1* apenas os itens mais relevantes. A exposição dos telespectadores aos espaços comerciais condiciona determinados indicadores, que também são fundamentais do ponto de vista do planeamento de *media*, como, por exemplo, as taxas de recordação das inserções e a capacidade dos mesmos para incentivar a aquisição dos produtos divulgados.

ÁREAS TEMÁTICAS DE INVESTIGAÇÃO NO ÂMBITO COMERCIAL	
Custos de publicidade e de programação	Linguagem e conteúdo na publicidade televisiva
Eficácia da publicidade e a sua produção	Questões éticas, regulatórias e legais
Estratégias de Ppublicidade	Implicações da nova tecnologia
Atitudes face à publicidade	Colocação de produtos
Exposição, atitudes e comportamentos dos telespectadores	Efeitos da programação na publicidade televisiva
TV por Cabo e por Satélite	Programas
Valores Culturais na publicidade televisiva	As inserções publicitárias: duração e efeitos decorrentes das posições
Vídeo, VOD e PVRs	<i>Zapping e Zipping</i> de espaços comerciais
Interactividade	Dimensão e composição da audiência
Internacional e países específicos	

Quadro F.7.1.1 - Áreas temáticas no âmbito da investigação das audiências comerciais

As publicações identificadas na área das análises presenciais e recorrendo aos dados de audimetria são uma vez mais escassas, apesar da sua relevância uma vez que permitem, assim como no âmbito programático, caracterizar em maior profundidade e com uma maior fiabilidade o comportamento dos telespectadores.

A investigação presencial é determinante para o planeamento de *media*: “Clearly these results do actually represent the underlying process of television audience behaviour at breaks, then the current views in presence research and media analysis will require some rethinking. No longer can presence be ignored in the optimal scheduling of advertising.” *Bunn (1982)*.

Apenas foi possível identificar um trabalho de previsão de audiências comerciais e um número muito reduzido de publicações que avalia a repercussão de determinadas variáveis estruturais na evolução dos índices de audiências, recorrendo a bases de dados de audimetria internacionais e essencialmente a análises de Regressão Linear (*Danaher, 1995; Green, 2007; Hammer, Riebe & Kennedy, 2009; Meurs, 1998, Tejada 1995*).

Quanto à caracterização e avaliação do comportamento de *zapping* (Gill, 2000) dos telespectadores e a análises que tenham como intuito inter-relacionar a vertente programática com a comercial, os trabalhos identificados recorrem quase na sua totalidade a análises de cariz mais qualitativo ou a inquéritos. Apenas foi identificado o trabalho de *Danaher (1995)*, em que se avaliam algumas influências estruturais no decréscimo da audiência dos espaços comerciais, recorrendo a dados de audimetria Austrálios registados ao segundo, a análises de Regressão Linear e aos períodos horários de maiores audiências. O autor também salienta o rigor destas fontes de dados, as limitações das abordagens qualitativas e as divergências de conclusões quando se comparam os resultados alcançados com a análise dos dados dos Estados Unidos face aos dados Australianos:

We study television audience levels during commercial breaks by recalibrating the people-meter to produce second-by second ratings. ... The purpose of this article is to study television audience levels during commercial breaks. ... The people-meter integrates a device that records precisely what the television is doing and a remote control that records what each person in the household is doing vis a vis their television viewing. ... *Danaher and Beed (1993)* found that 92 percent of panellists push their button correctly. ... The study period was the week of November 17 to 23, 1992, a typical viewing period for New Zealand, unaffected by any special social, economic, political or sporting events. ... Multiple Regression analysis was conducted ... It would certainly be interesting to see a replication of this research in other countries. ... Advertising recall studies are typically laboratory situations where respondents are exposed to an ad they have never seen before. Compare this with a typical advertising campaign where an advertisement is aired many times, thereby increasing the likelihood or recall immensely. Therefore, single exposure ad recall is unrealistic in practice, giving such recall studies limited value. Secondly studies which ask respondents to recall their typical behaviour during ad breaks produce results that might be shaped by societal opinion on advertising.

Segundo *Danaher (1995)*, as abordagens que se baseiam na observação através de situações experimentais, não apresentam qualquer validade uma vez que não correspondem a situações reais que permitam avaliar como a audiência se relaciona com o *media* e de acordo com o autor verificam-se as mesmas limitações quando se pretende conhecer as impressões e opiniões do público.

A quantificação das audiências comerciais apresenta uma grande relevância e a audimetria corresponde ao instrumento que o poderá viabilizar com u maior rigor (*Tejada, 1995*): “The Introduction and progressive development of Audience Rating makes available to professionals and efficient system for measurement and analysis of the audience for TV programs and commercial breaks, and thus informs decision-making.”

Segundo *Danaher (1995)*, também observamos de país para país, uma forte divergência nos resultados: “Zapping is likely to vary between countries according to the density of commercial breaks allowed and the number of alternative channels, which might have something more exciting to offer.” Consequentemente, são necessárias análises ajustadas à realidade nacional. Nas análises a realizar recorre-se à tipologia de dados que é considerada como sendo a mais adequada para a caracterização da exposição dos telespectadores aos espaços comerciais, assim como a metodologias que não constam da bibliografia identificada a nível internacional.

F.7.2. Objectivos

As análises a realizar às audiências comerciais inserem-se na denominada investigação presencial e visam o desenvolvimento de análises explicativas e, eventualmente, também com alguma capacidade de previsão das audiências comerciais quantificadas através do indicador *rating*. A sustentabilidade da indústria televisiva é função da vertente publicitária, consequentemente a previsão das audiências comerciais constitui um factor determinante para os vários agentes intervenientes neste sector de actividade. Para além da avaliação da repercussão de determinadas variáveis intrínsecas ao planeamento publicitário na evolução dos índices comerciais, a vertente programática também não deve ser negligenciada. As audiências comerciais, entre outros factores, também devem ser função dos níveis de exposição aos programas a que estão associados. Os restantes factores a considerar no desenvolvimento das análises explicativas dizem respeito a características de planeamento comercial, aos factores temporais, à contra-programação, notoriedade do canal e à estruturação da grelha, entre outras.

Pretende-se caracterizar a evolução dos índices comerciais das inserções publicitárias e, também, utilizando como dados simbólicos os espaços comerciais na sua totalidade, avaliando-se quais são os factores estruturais responsáveis pela variabilidade observada. As audiências das inserções publicitárias vão ser analisadas segundo duas vertentes, nomeadamente utilizando as taxas médias de audiências, mas, também, recorrendo aos seus valores relativos face ao desempenho dos programas a que estão associadas. Neste último contexto a variável será designada de *Spot_Ratio%* e resulta do quociente entre os valores de *ratings* comerciais e os valores de *ratings* programáticos. Avaliam-se as influências estruturais na evolução destes dois indicadores, assim como se é viável a sua previsão, ou em que medida os factores que condicionam em maior grau a sua evolução constituem apenas variáveis explicativas ou não. Por último, diferenciam-se os vários segmentos sociodemográficos relativamente à sua exposição aos espaços comerciais, considerando a evolução das respectivas taxas médias de audiências comerciais. As questões em avaliação correspondem às seguintes:

- Como influenciam os factores de planeamento a evolução das audiências comerciais, considerando cada inserção publicitária, nomeadamente o canal, as faixas horárias, a localização do espaço comercial e a sua estruturação, entre muitos outros factores que incluem os índices e as tipologias programáticas? Qual a influência das interacções observadas entre os factores estruturais representados nas bases de audimetria? Qual o grau de variabilidade dos *ratings* comerciais?
- Será viável prever os valores de *ratings*, com reduzidas margens de erro? As análises não lineares permitem-nos alcançar melhores resultados do ponto de vista da previsão face às análises lineares? As abordagens simbólicas lineares viabilizam melhores resultados face à aplicação das metodologias lineares tradicionais? Qual a tipologia de variáveis simbólicas mais adequadas para a representação dos factores em análise? As abordagens metodológicas que produzem melhores resultados diferenciam-se em função dos canais e horizontes temporais em análise? Em que medida os indicadores programáticos condicionam a evolução das audiências comerciais?

- Qual a influência sociodemográfica na evolução das audiências comerciais? Os homens, os jovens, as áreas rurais, os lares com maiores rendimentos apresentam uma menor exposição comercial? Quais são as variáveis sociodemográficas que permitem discriminar em maior grau os comportamentos de exposição aos espaços comerciais?

F.7.3. Bases de dados

Na seguinte *Figura F.7.3.1* consta a descrição da base de dados utilizada para a realização das análises propostas.

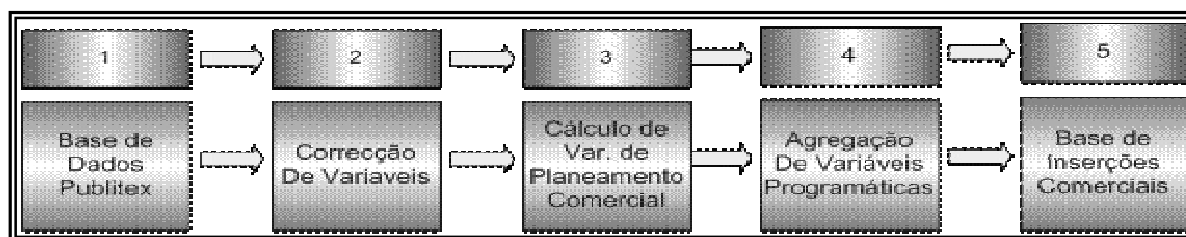


Fig. F.7.3.1 – Caracterização da construção das bases de dados de audiências comerciais

A base de dados corresponde aos índices de audiências que constam do *software* Publitex, a qual foi complementada com variáveis que visam inter-relacionar a componente programática com a comercial; que visam corrigir algumas das variáveis iniciais ou que caracterizam em maior profundidade algumas características intrínsecas ao planeamento publicitário. Os índices de audiências podem-se referir à totalidade do universo televisivo ou a determinados segmentos sociodemográficos, sendo as variáveis de segmentação coincidentes com as que foram referidas no contexto programático.

O período em análise coincide com o que foi adoptado nas análises programáticas por forma a atenuar a variabilidade decorrente da sazonalidade que caracteriza o consumo televisivo. As unidades estatísticas correspondem a inserções publicitárias ou a espaços comerciais transmitidos no decorrer de toda a semana a partir das 13H e apenas nos canais generalistas. Correspondem a 50.800 observações, das quais 11.700 se referem a inserções comerciais transmitidas nos canais estatais.

Foi necessário realizar algumas alterações às bases de dados Publitex, nomeadamente proceder ao cálculo de inúmeras variáveis, atendendo às incorrecções existentes na base de dados original, para além da ausência de informação relevante para as análises a desenvolver. Algumas das alterações realizadas foram as que se descrevem nos seguintes parágrafos:

(1) Na base de dados original associa-se o valor zero à posição de uma inserção comercial sempre que surge uma promoção do canal ou programática num espaço comercial. Esta tipologia de inserções publicitárias, também, não é contabilizada para o somatório do total de inserções por espaço comercial, sendo necessário recalcular este valor e a variável será designada por *Num_Inserções* e as correspondentes posições designadas por *Pos_Spot*. As posições são identificadas através de uma variável nominal com os seguintes níveis: primeira, segunda, intermédias, antepenúltima, penúltima e última, assim como, recorrendo a uma segunda variável do tipo ordinal.

(2) Calcular as variáveis que dizem respeito às promoções de programas ou canal, como, por exemplo, o número de autopromoções num espaço comercial (designada por *Num_Promos*), se a inserção corresponde a uma promoção (designada por *É_Promo*) e se uma inserção é precedida por esta tipologia de inserções (*Preced_Promo*), assim como, determinados cruzamentos e associações entre as variáveis mencionadas para viabilizar e enriquecer as análises recorrendo a algoritmos de regressão não linear de *Data Mining* proposicional.

(3) Calcular as durações globais desde o início do programa ou do espaço comercial até à transmissão da inserção (designadas por *Dur_BS* e *Dur_Ac* respectivamente) a nível absoluto e relativo, ou seja, a nível relativo face à duração global do programa e do espaço comercial, assim como, a duração do espaço comercial, a duração global de espaços comerciais transmitidos no decorrer de um programa e antes do seu início, os índices de saturação publicitária, entre outras, que pretendem caracterizar em maior profundidade os espaços comerciais. Calcular a localização do espaço comercial, variável que será designada de *Loc_Espaço_Comercial* que apresenta as categorias entre programas ou no seu decorrer e neste último contexto também o seu número, como por exemplo, 1º Intervalo até n.º intervalo.

(4) Calcular todas as variáveis que caracterizam os programas associados aos espaços comerciais ou aos programas adjacentes e concorrentes, assim como os índices programáticos, e não só, associados à envolvente do programa em que está inserido o espaço comercial em análise, uma vez mais a nível absoluto e relativo.

Os Modelos de Exposição Necessários à Aquisição de Publicidade no Sector Televisivo

Designação	Descrição	Designação	Descrição
Rat_Spot	Rating de uma inserção publicitária para a totalidade do painel	Ratio_Ldin%	Razão entre o rating de um programa e o rating do programa precedente
Dur_Spot	Duração da inserção (em segundos)	Ratio_Ldout%	Razão entre o rating de um programa e o rating do programa sucessor
Dur_Ac	Duração acumulada desde o início do espaço comercial até à transmissão da inserção (em segundos)	Dur_Fim	Duração desde a inserção publicitária e o final do espaço comercial
Dur_espaco_comercial	Duração global do espaço comercial	Ratio_Dur_Fim	Razão entre a duração desde a inserção publicitária e o final do espaço comercial e a duração do espaço comercial
Ratio_Dur_Ac	Razão entre Dur_Ac e Dur_espaco_comercial	Canal	Canal de transmissão
Loc_Espaco_comercial	Posição do espaço comercial: no decorrer de um programa ou entre programas	Shr_Prog% (T/P)	Share de audiência do programa associado à inserção publicitária
Num_Inserções	Número de inserções no espaço comercial ao que pertence a inserção publicitária em análise	Fid_Prog% (T/P)	Fidelidade ao programa
Pos_Spot	Posição da inserção no espaço comercial (numérico)	Time_Prog	Hora de início do programa
Ex_Promo	Se existem promoções de programas no espaço comercial ao qual pertence a inserção publicitária em análise	End_Prog	Hora do final do programa
Fol_Promo	Se a inserção publicitária sucede uma promoção	Syn_Ldin (1,2,3)	Se coincide a tipologia programática do programa associado ao espaço comercial e o precedente (três níveis)
Tipo_1/2/3 (1,2,3)	Tipologia programática do programa associado à inserção publicitária (três níveis)	Syn_Lout (1,2,3)	Se coincidem as tipologias programáticas do programa associado ao espaço comercial e o seguinte (três níveis)
Dia	Dia de semana	Rat_Prog% (CP) (T/P)	Rating dos programas em contra- programação
Spot_Ratio %	Razão entre Rat_Spot% e o rating programático parcial precedente	Dur_Ldin (P/T)	Duração global ou parcial do programa precedente (P-parcial, T-Global)
Sat_Prog%	Razão entre a duração global de um espaço comercial e a duração global programática	Dur_Prog (T/P)	Duração do programa (P-parcial, T-Global)
Periodo	Consideram-se as seguintes Faixa horárias: almoço (13-14), início de tarde (14-16), final de tarde (16-17), pré-prime-time (19-20), prime time (20-22) e noite (22-24)	Dur_BS	Duração desde o início do programa e a inserção
Tvr_Prog% (P/T) (P-Parcial, T-Global)	Taxas médias de audiências para o total da TV do programa parcial ou global associado à inserção publicitária	Tempo_Spot	Hora de transmissão da inserção
Rat_Ldin% / Rat_Ldout% (P/T)	Rating do programa global ou parcial anterior (Ldin) ou posterior (Ldout)	Nº_Partes	Número de partes que integram o programa
Rat_Prog% (P/T)	Rating do programa associado com o espaço comercial, total e parcial	Num_Promos	Número de promoções de programa no espaço comercial
		É_Promo	Se uma inserção é uma promoção
		Preced_Promo	Se uma inserção é precedida por uma promoção
		Pos_Spot	Posição da inserção (nominal)
		S_FS	Semana e Fim de Semana

Quadro F.7.3.1 – Descrição da base de dados de audiências comerciais

(5) Calcular o indicador *Spot_Ratio%* equivalente ao seguinte quociente: $Rat\_Spot\% / Rat\_Prog\%$ (Parcial) com o intuito de inter-relacionar a vertente programática e comercial.

F.7.4. Metodologias

Para avaliar a repercussão de determinadas variáveis não métricas no indicador *Spot_Ratio%* ou em *Rat_Spot%* e a influência sociodemográfica na evolução dos *ratings* recorre-se a análises ANOVA para as variáveis nominais, Testes T para Médias para as variáveis binárias e aos equivalentes algoritmos não paramétricos *Man Whitney U* e *Kruskal Wallis*, quando os pressupostos necessários à sua aplicação são violados. Para as variáveis numéricas calcularam-se os coeficientes de correlação linear das variáveis com os indicadores *Spot_Ratio%* e *Rat_Spot%*.

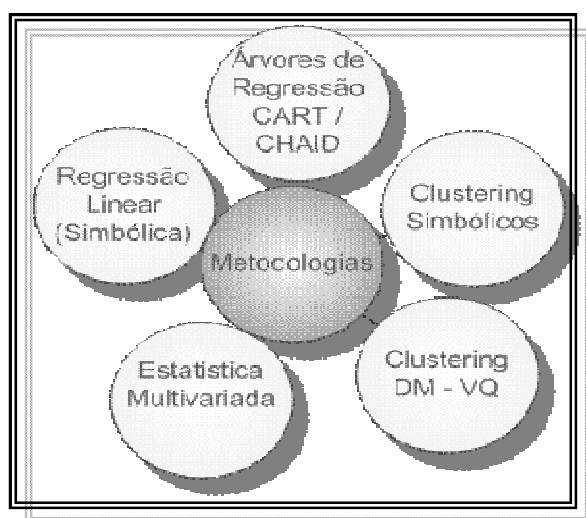


Fig. F.7.4.1 – Metodologias a utilizar nas análises comerciais

Nas análises Anova a I ou mais factores aplicaram-se os algoritmos que visam a análise de amostras com desigual dimensão ou com ausência de observações para determinados tratamentos. Aplicaram-se também determinados algoritmos de *clustering* e classificação simbólicos para a análise de espaços comerciais de forma agregada, atendendo à média de valores de *ratings* e à correspondente variabilidade.

Uma vez que se pretende o desenvolvimento de modelos explicativos de audiências comerciais, as metodologias de Regressão Linear ou as metodologias Não Lineares são as mais relevantes, recorrendo-se à Estatística Multivariada, ao *Data Mining* e ao *Knowledge Mining*, isto é, à Análise Simbólica. As Árvores de Regressão permitem-nos identificar as interacções existentes entre as várias variáveis explicativas consideradas e realizar uma previsão não linear. A Regressão linear Múltipla clássica e simbólica permitirá desenvolver modelos explicativos lineares de *ratings* comerciais e de valores de *Spot_Ratio%*.

As unidades estatísticas diferem consoante nos referimos a inserções publicitárias ou a espaços comerciais.

- Atendendo às análises que visam os indicadores associados aos espaços comerciais estes

Espaço comercial
associados a determinado
Programa, Dia e Localização

podem-se agregar em função dos programas, dias e localizações a que correspondem, por exemplo, `Telejornal_Terça_Espaço_Comercial_Antecede`, sendo caracterizados por variáveis simbólicas ou

manterem-se como entidades isoladas onde podem surgir algumas variáveis modais simultâneamente com as clássicas, como sejam, o período horário das inserções que o constituem, as tipologias ou as localizações de inserções e a caracterização dos produtos divulgados.

- Atendendo aos modelos explicativos simbólicos para as análises que visam as inserções publicitárias, os objectos de segunda ordem correspondem às inserções associadas a

Inserção associada a determinado
Programa, Dia, Localização
de Espaço comercial e
Posição da Inserção

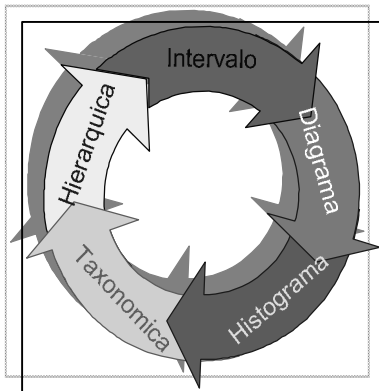
determinado Programa. Dia de semana, `Localização_Espaço_Comercial`, `Posição_Espaço_Comercial` ou a subconjuntos destas classificações, como, por exemplo, `Telejornal_Domingo_1º Espaço comercial_Última Posição`.

Como localização do espaço comercial considera-se se é anterior ao início do programa ou inserido no seu decorrer, identificando neste último caso o número do espaço comercial. Para a posição das inserções publicitárias considera-se a primeira, segunda, antepenúltima, penúltima, última e agregam-se todas as restantes no caso de existirem designando-as como posições intermédias. Esta classificação pretende diferenciar os resultados com base nas posições mais relevantes do ponto de vista comercial. Consideramos o dia de semana através de uma variável binária com as categorias semana e fim-de-semana para distinguir os vários episódios de um programa, uma vez que estes últimos variam nos seus horários de transmissão, o que irá condicionar as taxas médias de audiências para o total da TV e todos os restantes índices de audiência.

As variáveis explicativas consideradas nas análises dizem respeito aos índices comerciais que constam da base de dados associada ao *software* Publitex e aos índices programáticos que constam da base de dados associada ao *software* Telereport, assim como as variáveis temporais e de caracterização da estruturação da grelha e dos espaços publicitários (Quadro F.7.3.1).

Relativamente às variáveis dependentes neste capítulo pretende-se identificar as variáveis que apresentam uma maior repercussão na evolução dos *ratings* comerciais (variável *Rat_Spot%*) e também no indicador *Spot_Ratio%*, isto é, $Rat\_Spot\% / Rat\_Prog (P)\%$, que quantifica a inter-relação existente entre as audiências comerciais e os valores programáticos. Interessa avaliar a variabilidade observada não só nos seus valores absolutos, mas, também, nos valores relativos dos índices comerciais, uma vez que a valores mais elevados de *Rat_Spot%* poderá corresponder valores reduzidos da variável *Spot_Ratio%*.

As variáveis simbólicas são essencialmente da tipologia intervalo $[\alpha, \beta]$, nomeadamente as que resultam das variáveis numéricas nas bases de dados originais referentes a durações, índices de audiências programáticos e comerciais, entre outras.



Se as variáveis explicativas forem da forma $X_i = ([c_{i1}, d_{i1}], \dots, [c_{ik}, d_{ik}])$ e a dependente representada por $Y_i = [a_i, b_i]$, os elementos da matriz de variância e covariância V será: $V = (\text{cov}(x_{j_1}, x_{j_2})), j_1, j_2 = 1, \dots, k$

Fig. F.7.4.2 – Tipologias de variáveis simbólicas

para $j_1 \neq j_2$:

$$\text{cov}(X_{j_1}, X_{j_2}) = \frac{1}{4n} \sum_{i=1}^n (c_{ij_1} + d_{ij_1})(c_{ij_2} + d_{ij_2}) - \frac{1}{4n} \left[\sum_{i=1}^n (c_{ij_1} + d_{ij_1}) \right] \left[\sum_{i=1}^n (c_{ij_2} + d_{ij_2}) \right]$$

e para $j_1 = j_2$:

$$\text{var}(X_{j_1}) = \frac{1}{4n} \sum_{i=1}^n (c_{ij_1} + d_{ij_1})^2 - \frac{1}{4n} \left[\sum_{i=1}^n (c_{ij_1} + d_{ij_1}) \right]^2$$

Os coeficientes de Regressão β calculam-se por $(X'X)^{-1}X'Y$, sendo o elemento $(j+1)$ do vector $(K+1) \times 1$ de $X'Y$ dado por $(x'_j y_j) = \frac{1}{2} \sum_{i=1}^m (c_{ij} + d_{ij})(a_i + b_i)$, $j=1, \dots, k$.

Como variável nominal consta das bases de dados o canal de transmissão de cada inserção publicitária. Cada modalidade da variável nominal converte-se numa variável binária e retiramos um nível, assim como quando se utilizam as variáveis *dummy* na Regressão Linear não simbólica. Sendo K o número de variáveis e se V_{im} representar o valor da modalidade m da variável I então $\forall i, j$ de 1 a K desde que $\sum_m V_{im} = \sum_m V_{jm} = 1$ então $(X'X)$ será não reversível.

Também constam das bases de dados variáveis da tipologia diagrama, como, por exemplo, os períodos horários associados à transmissão de uma inserção ou a posição da inserção, uma vez que os programas alteram a sua posição na grelha de programas ao longo da semana. Por exemplo: (a) (0.6) Horário Nobre, (0.4) Noite, significando que em 60% das observações a inserção em análise foi transmitida no período do horário nobre; (b) as sinergias tipológicas face a programas adjacentes ou em contra-programação (ex. (0.3) Syn_Ldin=TRUE; (0.7) Syn=False, significando que para 30% das observações as tipologias programáticas coincidem ou (c) a distribuição das tipologias programáticas em contra-programação ou adjacentes (ex. (0.2) telenovela; (0.6) espaços informativos; (0.2) desporto).

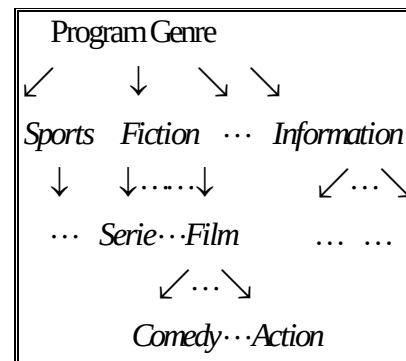
<i>Sports</i>	<i>News</i>	<i>Serie</i>	<i>Film</i>
0,2	0,3	0,4	0,1
0	0	0,9	0,1
0,4	0,5	0	0,1

Podemos, também, considerar a distribuição de categorias de produtos divulgados, a distribuição das tipologias e as localizações de inserções nos espaços comerciais associados às inserções publicitárias em análise. Para estas variáveis define-se uma matriz com uma coluna para cada modalidade da variável diagrama e cada célula representa o peso da modalidade a que se refere a coluna.

Finalmente, como variáveis histograma podemos considerar a distribuição das durações das inserções num espaço comercial (ex. $(0,7[15,30],0,3[30,60])$, ou seja, 70% dos espaços comerciais apresentam durações entre os 15 e os 30 segundos) ou a distribuição dos *ratings* de programas em contra-programação.

As variáveis histograma X_j para uma observação i representam-se por $X_{ij} = [p_{ij1}[c_{ij1}, d_{ij1}], \dots, p_{ijs_{ij}}[c_{ijs_{ij}}, d_{ijs_{ij}}]]$ onde $p_{ijs_{q1qj}}$ será a frequência relativa do subintervalo $[c_{ijs_{q1qj}}, d_{ijs_{q1qj}}]$, $S_{q1qj} = 1, \dots, s_{ij}$, $i=1, \dots, n$, $j=1, \dots, K$. As variáveis da tipologia intervalo podem ser substituídas por variáveis histograma se necessário.

A variável que representa a tipologia programática será da tipologia taxonómica, apresentando três níveis. Em cada nível Z existe um número diferente de categorias (por exemplo, Ficção (L1), Filme (L2), Aventura (L3) que descrevem o género televisivo de um filme de aventura). Vamos utilizar o método de agregação para realizar a Regressão desta tipologia de variáveis (Bock & Diday, 2000).



Se uma observação apresentar um valor no nível j , os seus valores descendentes serão agregados para esse nível. Obtemos, assim, uma única regressão através dos diversos níveis e todas as observações são utilizadas na análise. Apesar do método alterar a matriz inicial dos dados, os valores permanecem fieis à bases de dados real e obtemos uma só regressão que inclui todos os níveis da variável. Nas bases de audimetria existem inúmeros níveis nos níveis mais elevados, logo não se agrega excessivamente as observações. A posição de um espaço comercial ou o dia de transmissão também podem ser codificados como variáveis taxonómicas.

Nas análises não lineares a realizar recorreremos ao algoritmo CART e para nos ajudar na interpretação das Árvores de Regressão CART, *Breiman et al. (1984)* sugerem a medida M representada pela seguinte equação (1) que traduz a relevância global das variáveis explicativas utilizadas.

A definição de M baseia-se na redução da diversidade que resulta da selecção da variável X_j como *splitter*, ou pela sua potencial utilização como variável substituta. Se existir mais de uma subdivisão substituta do nó O associado com X_j , escolhemos a que maximiza $\Delta I(\Pi^j, O, A)$.

$$M(X_j) = \sum_{O \in A} z^{jO} \Delta I(\Pi^j, O, A) \quad (\text{Eq. 1})$$

$$z^{jO} = \begin{cases} 1, & \Pi^j \in \Pi^s(O) \\ 0, & \Pi^j \notin \Pi^s(O) \end{cases} \quad \begin{array}{l} \text{onde, } \Pi^j \text{ corresponde ao } \textit{split} \text{ do nó } O \text{ associado com a} \\ \text{variável explicativa } X_j \text{ e } \Pi^s \text{ corresponde ao grupo de} \\ \textit{surrogates}, \text{ e } \Delta I(\Pi^j, O, A) \text{ corresponde ao decréscimo na} \\ \text{diversidade da Árvore global (A) como consequência de} \\ \text{utilizar o } \textit{splitter } X_j \text{ no nó } O. \end{array}$$

Se os *splits* forem univariados as variáveis podem ser ordenadas numa escala de 0 a 100, de acordo com o seu potencial de poder explicativo: $M(X_j) = 100 * [M(X_j) / \max M(X_j)]$. Esta medida será comparada com os coeficientes β da Regressão Linear Múltipla para comparar a capacidade explicativa de cada uma das variáveis segundo ambas as metodologias. Contudo, nas Árvores, também, podemos utilizar uma combinação linear das variáveis explicativas numéricas que constam das bases de dados.

F.7.5. Resultados

Os resultados subdividem-se nas seguintes quatro partes:

I. Análises exploratórias em que se avalia o grau de associação existente entre as variáveis explicativas, que eventualmente também podem apresentar capacidade de previsão, e as variáveis dependentes. Também, constam os resultados das análises realizadas aos espaços comerciais de forma agregada;

II. Análises explicativas de Regressão linear e a posterior avaliação da sua capacidade de previsão;

III. Análises explicativas não lineares e a posterior avaliação da sua capacidade de previsão;

IV. Análises da influência sociodemográfica na evolução dos *ratings* comerciais.

I. Resultados das Análises Exploratórias

Os indicadores comerciais representam os dados mais relevantes para o planeamento de *media*. Segundo *Danaher (1995)*, os valores dos *ratings* decrescem no decorrer dos espaços comerciais, mas de forma pouco significativa. O valor médio observado no decorrer do horário nobre corresponde a um decréscimo de apenas 5%. O autor, também, conclui que as características dos espaços comerciais, como, por exemplo, o número de inserções publicitárias apresentam uma repercussão pouco significativa no valor das reduções observadas nos *ratings* comerciais. Salienta que existem poucas causas sistemáticas que sejam responsáveis pelas alterações na evolução do indicador *Spot_Ratio%*. Por conseguinte, as análises que constam deste capítulo pretendem avaliar qual é a capacidade explicativa e de previsão das variáveis que

constam das bases de dados de audimetria. Os resultados são apresentados em função dos canais e, por vezes, apenas para determinados períodos horários.

Ia. Alguns Resultados para o Canal RTP1

A base de dados está constituída por 11.713 observações, ou seja, inserções publicitárias repartidas por 752 espaços comerciais. Os resultados encontram-se subdivididos em função da variável dependente: *Spot_Ratio%* e *Rat_Spot.%*

➤ Grau de Associação existente entre determinadas variáveis nominais e a variável dependente *Spot_Ratio%*

(1) Avaliando a repercussão da posição de uma inserção num espaço comercial nos valores médios de *Spot_Ratio%* verificamos que apenas as duas posições iniciais e a final apresentam uma maior retenção dos níveis de audiência. As restantes posições são mais penalizadas, não se diferenciando entre si. A primeira posição apresenta o maior valor médio de *Spot_Ratio%*, no valor de 0.98. Um valor unitário corresponderia à inexistência de redução nos valores dos *ratings comerciais* face aos valores dos *ratings programáticos*. Segue-se a segunda posição (*Spot_Ratio_Médio*=0.94), a última posição (*Spot_Ratio_Médio*=0.86) e em último lugar, segundo os testes Post Hoc da *Análise Anova* realizada ($F(5,11707)$, p valor>00000), o conjunto constituído pelas posições intermédias, antepenúltimas e penúltimas.

Assim, apesar do canal estatal apresentar um perfil de audiências que em princípio terá uma menor predisposição para o desenvolvimento de comportamentos de “zapping”, verificamos que o decréscimo nos valores dos *ratings* comerciais a nível médio pode ser considerado significativo. A última posição representa em média 86% do valor da taxa de audiência programática. Uma vez que as inserções publicitárias têm uma reduzida duração, este decréscimo deve-se em maior grau ao diferencial no número de telespectadores, e, não, à duração da exposição.

(2) Os espaços comerciais localizados entre programas perdem menos audiência, o que poderá ser atribuído, em parte, aos efeitos de herança existentes entre programas adjacentes e ao planeamento publicitário que recorre a uma ligeira redução do número de inserções médio por espaço comercial para esta localização (0,87 de valor médio de *Spot_Ratio%* vs. 0,77 para os espaços comerciais no decorrer de um programa). O menor valor médio global de *Spot_Ratio%* para os espaços comerciais localizados no decorrer de um programa deve-se aos valores mais reduzidos que estão associados aos primeiros e segundos intervalos. Observamos alguma correspondência, mas inversa, entre as durações globais dos espaços comerciais ou o número de inserções e os valores médios de *Spot_Ratio%*. Segundo *Danaher (1995)*, uma das variáveis mais significativas corresponde ao número de inserções num espaço comercial, a qual se relaciona de forma inversa com este indicador. Os espaços comerciais com um maior número de inserções têm uma maior duração, o que origina uma maior perda de audiência.

Este autor não testa a localização do espaço comercial como variável explicativa deste comportamento.

(3) Verificando a interação existente entre a localização do espaço comercial e as posições das inserções publicitárias (*análise Anova a II factores*) na variável em estudo, verificamos que as posições intermédias e as finais dos primeiros e segundos espaços comerciais são as que contribuem para as suas menores audiências médias globais. Os restantes espaços comerciais apresentam uma média superior, uma vez que apresentam uma maior regularidade no seu desenrolar e os espaços comerciais que antecedem os programas são favorecidos pela permanência da audiência no seu início e no final, enquanto, esperam pelo início do programa sucessor. Os programas com um maior desempenho, por exemplo, os eventos desportivos conseguem manter um fluxo de audiências mais uniforme no decorrer dos seus espaços comerciais.

Apesar do canal RTP1 apresentar uma audiência com uma composição sociodemográfica, aparentemente menos susceptível a desenvolverem comportamentos de *zapping*, os seus espaços comerciais podem apresentar reduções que atingem 30% dos valores dos *ratings* programáticos. Esta diferenciação ao nível da evolução de *Spot_Ratio%* em função das localizações de um espaço comercial deverá estar relacionada com a sua duração. Logo, para os espaços comerciais localizados entre programas, que são mais breves apenas as posições intermédias contribuem em maior grau para a redução dos valores de *Spot_Ratio%*.

(4) Analisando a influência do período horário nos valores médios de *Spot_Ratio%* foi possível retirar as seguintes conclusões: As inserções da tarde (16H-19H), do princípio da tarde (13H-16H) e do horário nobre (20H-22H) são as que apresentam um maior valor médio de *Spot_Ratio%* (0.90). Seguem-se as inserções da noite (22H-24H) (0.80) e por último do pré-prime-time (19H-20H) e da madrugada (0.68) (24H-02H).

Danaher (1995) apenas considerou nas suas análises inserções do período do horário nobre, e que a nível nacional também apresentam menores reduções nos valores de *ratings*. Para o horizonte temporal em análise o valor médio observado é de 10% e segundo *Danaher*, na Austrália este valor corresponde a apenas 5%. Não se observam diferenças entre os períodos da tarde e o horário nobre. Apesar das audiências programáticas serem maximizadas no horário nobre, aos valores de *Spot_Ratio%* associamos intervalos de valores mais elevados. De acordo com *Danaher (1995)*, os valores dos *ratings* programáticos levam a valores mais elevados de *Spot_Ratio%*. Para o canal RTP1 apesar do decréscimo das audiências comerciais, verificamos que existe proporcionalidade entre a evolução dos *ratings* programáticos e os valores comerciais. As variáveis relacionam-se de forma proporcional.

Ambos os espaço informativos e o decorrer da tarde, até às 19H são os períodos onde se observam menores reduções nas audiências comerciais face aos valores programáticos. Corresponde aos períodos onde poderá existir uma audiência mais regularmente disponível e com uma maior incidência em telespectadores

que não exercem maioritariamente uma actividade laboral. No decorrer da noite onde se transmitem os programas sucessores ao espaço informativo, os telespectadores apresentam uma maior actividade durante os blocos de anúncios, assim como os programas que antecedem o espaço informativo da noite e os programas transmitidos no decorrer da madrugada. A RTP1 corresponde ao canal que é, por vezes, mais penalizado quanto aos índices de audiências dos programas de ficção da noite. Consequentemente, durante os seus blocos de anúncios o público deverá sintonizar os canais privados, ou realizar outras actividades. A madrugada apresenta não só menores valores de *ratings* programáticos, como também, uma maior queda nas audiências comerciais. Assim como se verificou para as análises de índices de lealdade programática, em horários mais tardios não existe obrigatoriamente uma audiência de menor dimensão, mas mais regular e mais passiva no decorrer dos espaços comerciais.

(5) Analisando a interacção existente entre os períodos horários e as posições das inserções, verificamos que as inserções da noite e da madrugada são penalizadas por apenas as duas posições iniciais dos espaços comerciais apresentarem maiores valores de *Spot_Ratio%* e os espaços comerciais do pré-prime-time por se verificar o inverso. O horário nobre é favorecido pela exposição da audiência ao início e ao final dos espaços comerciais. Os espaços comerciais da tarde são favorecidos por um comportamento mais regular da audiência no seu decorrer e o princípio da tarde por ambos os aspectos.

Uma vez que o perfil das audiências se diferencia no decorrer do dia, ainda que de forma menos significativa face ao que se verifica nos canais privados, inevitavelmente os horizontes temporais influenciam a evolução dos *ratings* no decorrer dos espaços comerciais. Atendendo aos maiores níveis de exposição aos espaços informativos da noite, existe uma menor queda das audiências nas posições iniciais e finais. Os espaços comerciais do horário nobre são favorecidos nas suas posições finais por uma maior proximidade face ao início do espaço informativo. A tarde acaba por apresentar um comportamento mais regular da audiência.

(6) A nível médio, não podemos afirmar que exista uma menor redução dos *ratings* comerciais durante o fim-de-semana face ao que se verifica em dias laborais. Os vários dias coincidem em que os espaços comerciais antes do início dos programas saem beneficiados por uma maior retenção da audiência nas posições iniciais e finais e para as restantes localizações apenas nas posições iniciais. Pontualmente, verifica-se alguma tendência para a regularidade no seu desenrolar. Interessa sobretudo realçar que a redução ao nível dos índices de audiências das posições intermédias é superior ao fim-de-semana. Apesar da maior disponibilidade do público, existe uma maior actividade no decorrer dos espaços comerciais o que poderá ser atribuído a um público com uma composição sociodemográfica diferente, onde não existe um peso tão significativo das maiores faixas etárias.

(7) Considerando as tipologias programáticas de Grau 2, verificamos que apresentam valores de *Spot_Ratio%* superiores à unidade: os documentários, programas de política, reportagens, programas de variedades, filmes western e sketches humorísticos.

Segundo *Danaher (1995)*, as tipologias influenciam a evolução de *Spot_Ratio%*, mas os resultados estão essencialmente influenciados pelas durações que as diversas tipologias programáticas apresentam. Para este canal, este factor poderá ter alguma relevância. Contudo, a maior fidelidade que é dirigida aos programas informativos também poderá condicionar os resultados.

(8) Atendendo ao papel desempenhado pelas promoções, verificamos que não existe uma diferença estatisticamente significativa entre a média de *Spot_Ratio%* associada às inserções que são, ou não, precedidas por promoções, ou entre as inserções que correspondem, ou não, a promoções do canal e programáticas. *Danaher (1995)* apenas avalia a influência da existência de uma promoção no início ou no final do espaço comercial e conclui que a sua existência leva a maiores valores de *Spot_Ratio%*. Perante o horizonte temporal em análise esta tipologia de inserções publicitárias não influencia os resultados.

➤ **Grau de Associação existente entre determinadas variáveis nominais e a variável dependente *Rat_Spot%***

Realizaram-se idênticas análises considerando como variável dependente as taxas médias de audiências para as inserções publicitárias.

(1) As primeiras, segundas e últimas posições apresentam maiores valores de *ratings* (*Rat_Spot_Médio*=5,3%). As intermédias, antepenúltimas e penúltimas têm os menores valores (*Rat_Spot_Médio*=4,6%), não se diferenciando entre si. No entanto, é reduzida a diferença entre os dois conjuntos de posições, com um valor de apenas 0,7%, a nível médio.

(2) Relativamente à localização do espaço comercial, os primeiros intervalos e os intervalos que antecedem os programas são os mais beneficiados. Os espaços comerciais associados a programas mais longos com um maior número de intervalos, também são transmitidos nos períodos de menores audiências programáticas e comerciais.

(3) Cruzando a posição e a localização de uma inserção verificamos que as primeiras e as segundas posições destacam-se em especial para os intervalos que se localizam antes do início dos programas e para os primeiros intervalos. Os segundos e os terceiros intervalos também apresentam um acréscimo de valores nas posições iniciais face às intermédias, mas menor. As últimas posições sobressaem para os espaços comerciais antes do início dos programas e para os terceiros. Para os quartos intervalos até aos sextos verifica-se uma maior regularidade dos *ratings*. Portanto, uma vez que a maioria dos intervalos estão associados à primeira parte dos programas, apenas as duas primeiras posições e a última se destacam, observando-se uma redução muito significativa ao nível das posições intermédias.

(4) Considerando a repercussão das faixas horárias nos valores dos *ratings*, verificamos que as inserções da madrugada apresentam os menores valores de *ratings* comerciais, segue-se a tarde, a noite, o princípio da tarde, o pré-prime-time e por último o prime time corresponde ao período que maximiza esta variável. As audiências comerciais maximizam-se para os espaços informativos e para os seus programas adjacentes. A noite também apresenta maiores valores em relação à madrugada.

(5) Em relação às interações entre o factor período horário e as posições das inserções, os programas de prime time apresentam uma evolução mais variável, contrastando com a noite onde se verifica uma certa regularidade. Os períodos de maiores audiências também apresentam uma maior variabilidade dos *ratings*. Durante a tarde, a noite e a madrugada existe uma tendência para a regularidade de valores de *ratings* comerciais independentemente das localizações dos intervalos.

No pré-prime-time assistimos a uma redução nos valores dos *ratings* associados às posições intermédias independentemente da localização do espaço comercial. No prime time para os primeiros e segundos intervalos destacam-se as posições iniciais. Para os espaços comerciais exibidos antes do início dos programas as posições intermédias são as mais penalizadas e para os terceiros destacam-se com maiores valores de *ratings* as primeiras e as últimas posições.

(6) O fim-de-semana apresenta uma média de *ratings* comerciais equiparável aos valores dos dias laborais. Apesar de existir uma audiência média para o total da TV de maior dimensão, a nível médio não se observam diferenças entre os *ratings* comerciais para cada um destes períodos.

(7) As tipologias que apresentam maiores valores de *ratings* comerciais correspondem aos espaços informativos fa noite, sociedade, política, musicais e os sketch humorísticos com valores no intervalo [7,9%;9,6%]. Seguem-se o futebol, documentários, programas de conhecimento e de automobilismo com valores no intervalo [5,1%;7,2%]. Com menores audiências comerciais destacam-se a maior parte dos filmes das tipologias: western, terror, acção, drama, aventura, romance, ficção científica e suspense com valores no intervalo [0,7%; 2,2%]. Os filmes também são transmitidos em faixas horárias mais tardias. Os espaços informativos e os programas sucessores maximizam estes valores.

(8) As promoções de programas não apresentam maiores valores de *ratings* relativamente às restantes tipologias e a precedência de uma inserção por uma promoção também não influencia esta variável. Calculando a média de *Rat_Spot%* em função do número de promoções, os espaços comerciais com um número superior a cinco são os que apresentam maior média da variável em estudo.

➤ Análises a Valores de Índices de Audiências de Espaços Comerciais

Antes da avaliação da linearidade das relações existentes entre as variáveis explicativas e *Rat_Spot%* ou *Spot_Ratio%* procedemos à análise dos espaços comerciais de forma agregada. Associa-se a cada espaço comercial que integra as bases de dados o valor médio dos *ratings* comerciais e o diferencial de valores. Para os dados agregados e simbólicos consideram-se os espaços comerciais associados a determinado programa, localização e dia, por exemplo, 1º Espaço comercial_espaço_informativo_noite_FS.

Pretende-se identificar quais são os factores responsáveis em maior grau pela variabilidade de *ratings* comerciais. Para o efeito utilizar-se-á como algoritmo de *clustering* de *Data Mining* o Kohonen VQ (“Vector Quantization”), implementado no *software* SAS Enterprise Miner. Este algoritmo corresponde a uma Rede Neuronal não supervisionada, com um funcionamento muito semelhante ao do algoritmo de *clustering* de K Média. Em alternativa aplicou-se o algoritmo de *clustering* simbólico divisivo a dados agregados simbólicos. Realiza-se o *clustering* dos espaços comerciais em função dos valores médios e de variabilidade de *ratings* por forma a tentar identificar se alguns factores podem estar associados a maiores médias de *ratings* comerciais, mas, também, a uma maior estabilidade dos seus valores.

(1) O *clustering* VQ com base nos valores de *ratings* e respectivo diferencial de valores permitiu identificar que os seus valores são essencialmente função dos períodos horários, tipologias e das localizações dos espaços comerciais.

Os maiores valores médios dos *ratings*, mas simultaneamente mais variáveis com uma média de 23,7% e com um diferencial de valores de 5,5% correspondem a alguns espaços comerciais antes do espaço informativo da noite e aos primeiros intervalos de alguns eventos desportivos, transmitidos no pré-prime-time e no prime time.

Os espaços comerciais que apresentam maiores valores, mas menos variáveis com uma média de 10,8% e um diferencial de 1,46% correspondem aos períodos de pré-prime-time, prime time e da noite e às tipologias de informação, desporto e divertimento. Existe uma maior incidência em localizações entre programas no pré-prime-time e noite e no prime time também no decorrer do espaço informativo da noite.

Com valores mais reduzidos e menos variáveis com uma média de 4,2% e um diferencial de 1,16% ou uma média de 7,7% e um diferencial de 0,9% existe uma grande amplitude de períodos horários e de tipologias, sendo difícil identificar alguma variável com uma distribuição mais homogênea que os permita caracterizar. Não podemos afirmar que os menores valores de *ratings* estejam restringidos a determinados períodos ou programas. Em todos os períodos identificamos comportamentos muito variáveis da audiência no decorrer dos espaços comerciais. Foi mais fácil identificar os espaços comerciais que maximizam as audiências, correspondendo aos períodos de maiores audiências e aos programas mais populares, ou seja, ao espaço informativo da noite e a determinados jogos de futebol.

Os eventos desportivos não só permitem a maximização das audiências comerciais e programáticas, como também, um significativo acréscimo na variabilidade dos valores dos *ratings* comerciais.

Segundo *Danaher (1995)*, existem poucas causas sistemáticas que possam provocar o diferencial de valores de *ratings* dos espaços comerciais. Evitam-se os espaços comerciais de forma aleatória e não existem características individuais ou do lar que se correlacionem com os comportamentos de *zapping*. Assim como, no campo programático alguns autores defendem que as características sociodemográficas não se associam com padrões diferenciados de exposição televisiva, também a nível comercial este autor defende a fraca repercussão destas variáveis. No último capítulo de apresentação de resultados consta a influência sociodemográfica nos valores dos *ratings* comerciais.

(2) Por aplicação do algoritmo DIV a dados simbólicos verificamos que os maiores valores de *ratings* correspondem apenas a determinados jogos de futebol e ao espaço informativo da noite no decorrer de toda a semana. Quanto mais relevante for o evento desportivo menor será o diferencial de valores de *ratings* no decorrer dos intervalos. Os programas mais favorecidos neste período envolvem o FC Porto, o Sporting e o Benfica. Com valores intermédios entre os 7% e os 11% também correspondem a alguns jogos de futebol, ao programa de sucessor do espaço informativo da noite, ao Festival da Eurovisão, ao espaço informativo da tarde e a alguns eventos como a Festa de 1ª de Maio.

Ib. Resultados para o Canal TVI

A TVI diferencia-se da RTP1 por ser um canal mais comercial, proporcionando aos telespectadores uma grelha de programas que tenta corresponder às preferências dos mesmos. Para o período em análise apresenta uma tendência crescente ao nível dos seus *shares* de audiências.

➤ Grau de Associação existente entre determinadas variáveis nominais e a variável Dependente *Spot_Ratio%*

Realizando análises Anova ou Teste T para diferenças de médias foi possível retirar as seguintes conclusões:

(1) Os *ratings* comerciais dos espaços comerciais localizados entre os programas superam os valores dos intervalos situados no seu decorrer, em parte devido aos efeitos de herança e que são favorecidos pela sua menor durabilidade total. Coincide com os resultados do canal RTP1.

(2) As localizações apresentam a mesma margem diferencial anteriormente calculada para a variável *Spot_Ratio%* do canal RTP1 (0,1 de diferencial) e portanto é reduzida. Diferenciando as diversas localizações, os segundos intervalos minimizam os valores de *Spot_Ratio%* (0,74), seguem-se os primeiros conjuntamente com os terceiros (0,83), os intervalos que antecedem os programas com os quartos (0,92) e em primeiro lugar o quinto intervalo com um valor médio de 0,99.

As maiores médias de valores de *Spot_Ratio%* estão associadas aos últimos espaços comerciais de programas mais longos, nomeadamente os quartos e quintos, ou seja, os menos representativos, uma vez que constituem apenas 2,1% da base de dados e aos espaços comerciais que antecedem os programas.

(2) As duas posições iniciais de um espaço comercial são as que apresentam uma maior média (1,05). As restantes posições não se diferenciam entre si, com uma média no valor de 0,85. Para este canal a nível médio as últimas posições não apresentam valores mais elevados.

(3) Cruzando a posição e a localização do espaço comercial (*Análise Anova a II factores*) verificamos que independentemente da localização do espaço comercial as duas primeiras posições são as que assumem maiores valores e as restantes não se diferenciam entre si. Portanto, para a TVI a última posição não se destaca em nenhum contexto, contrastando com os resultados da RTP1.

(4) Os períodos horários com um menor valor médio de *Spot_Ratio%* são o pré-prime-time (0,77), seguindo-lhe o princípio da tarde (0,81), o prime time agregado à noite (0,85), a madrugada (0,89) e com maiores valores a tarde (0,93). Estes resultados coincidem com os apresentados para o canal RTP1 apenas para os períodos de pré-prime-time e da tarde, uma vez que a madrugada era um dos períodos com valores mais reduzidos e o prime time e a noite apresentam maiores valores.

Para a TVI o período com um menor decréscimo das audiências comerciais corresponde à tarde e à madrugada e segue-se o prime time e a noite. Correspondem aos períodos que também maximizam as audiências programáticas para o horizonte temporal em análise. Logo, as maiores faixas horárias podem apresentar uma maior estabilidade na queda dos valores das audiências, contrastando com o que se verifica para a RTP1. Contudo, a grelha da RTP1 apresenta a transmissão de inúmeros filmes enquanto que a TVI não incide apenas nesta tipologia. As tipologias e as faixas horárias por si só não garantem que se alcancem determinados resultados, os quais também se encontram dependentes dos canais, entre outros factores.

(5) Para o princípio da tarde, prime time e noite independentemente das localizações dos espaços comerciais destacam-se as duas posições iniciais com os maiores valores de *Spot_Ratio%*. No decorrer do principio da tarde apenas as duas primeiras posições dos intervalos que antecedem os programas se diferenciam das primeiras posições dos restantes espaços comerciais. Para a noite pode-se estabelecer a seguinte ordem decrescente de valores de *Spot_Ratio%*: espaço comercial que antecede os programas, primeiros e segundos intervalos para todas as posições. No prime time, os primeiros intervalos apresentam uma maior média para todas as posições. Os espaços transmitidos antes dos programas e os segundos intervalos apenas se diferenciam para as duas posições iniciais, sendo superiores para os segundos intervalos.

Na madrugada existe uma redução muito acentuada de valores de *ratings* a partir da segunda posição exclusiva para todas as localizações. Os espaços comerciais antes dos programas apresentam uma maior média marginal. Para todas as posições estes espaços comerciais diferenciam-se significativamente face aos restantes. Para o primeiro intervalo, apenas as duas primeiras posições se diferenciam das restantes. A partir das posições intermédias não se observam diferenças entre os primeiros, segundos e os terceiros intervalos.

(6) Os fins-de-semana apresentam um valor significativamente mais reduzido de *Spot_Ratio%* (0,80 face aos 0,90 de valor médio semanal) observando-se os maiores valores à Quarta e à Quinta (0,94). As Segundas, Terças e Sextas não se diferenciam entre si apresentando um valor de 0,88. Ao contrário do que se verifica na RTP1 os dias laborais versus fim-de-semana diferenciam-se, com os dias laborais a apresentarem um menor decréscimo dos valores das audiências. Apesar da maior disponibilidade do público os espaços comerciais ao fim-de-semana são os mais penalizados.

(7) As inserções que não são promoções e que não são precedidas por esta tipologia de inserções apresentam uma média equiparável (0,87/ 0,84). Em relação ao número de promoções por intervalo, os espaços comerciais isentos de promoções ou com apenas uma apresentam um maior valor médio de *Spot_Ratio%*.

A partir de uma promoção vai decrescendo o seu valor até se alcançarem cinco ou seis promoções, a partir do qual onde se verifica um acréscimo no valor desta variável. Contrasta com os resultados da RTP1, segundo os quais a uma maior média estão associados os intervalos com mais de cinco promoções.

(8) Considerando as tipologias de grau 1, os programas com os maiores valores médios de *Spot_Ratio%* correspondem aos programas de juventude, cultura geral e conhecimento e de desporto. Para as tipologias de grau 2 destacam-se os programas de concursos, informação especializada, juvenis, magazines, programas religiosos, desporto e as transmissões desportivas. Para as tipologias de grau 3 os maiores valores estão associados a programas de conhecimento, bolsa, meteorologia, futebol e informação cultural (valores no intervalo [1,0;1,5]). Seguem-se com valores no intervalo [0,85;0,90] por ordem crescente de valores de *Spot_Ratio%* as tipologias de filmes de aventura, casos reais, revistas, misto, telejornais (0,88), e *talk shows*. Com valores no intervalo [0,70;0,90] constam os programas das tipologias de conhecimento, filmes épicos, fantasia, drama, comédia, acção. Os filmes têm os maiores valores de índices de saturação publicitária e simultaneamente os menores valores de *Spot_Ratio%*. Portanto, existe uma relação inversa entre as duas variáveis. Em relação às durações programáticas os programas com menores valores de índices de saturação publicitária também correspondem aos menores valores de durações globais e a maiores valores de *Spot_Ratio%*.

➤ **Grau de Associação existente entre determinadas variáveis nominais e a variável Dependente *Rat_Spot%***

(1) As inserções no decorrer de um programa apresentam os menores valores de *ratings* com um valor médio de 4,8% vs. 6,6% para as inserções entre programas. As inserções dos intervalos localizados antes do início dos programa apresentam um valor médio de *ratings* comerciais de 6,5%; as inserções dos primeiros intervalos um valor médio de 5,2%; os segundos, terceiros e os quintos intervalos de 3,7% e nos quartos intervalos de 2,8%. Não existe uma correspondência directa entre as durações e *Rat_Spot%*. Esta relação é um pouco mais explícita considerando a variável *Rat_Spot_Médio%* por espaço comercial, podendo-se associar para algumas das localizações, maiores valores aos *ratings* às menores durações dos espaços comerciais.

(2) A primeira posição de um espaço comercial apresenta um maior valor de *rating* comercial médio (7,1%) e a segunda também (6,9%). As posições intermédias, antepenúltimas e últimas não se diferenciam entre si, com uma média no valor de 5,8%.

(3) Verificamos que para os espaços comerciais que antecedem os programas as duas primeiras posições apresentam maiores valores de *ratings*. Estes índices reduzem-se nas posições intermédias e as três posições finais ainda que com valores inferiores aos iniciais apresentam um ligeiro acréscimo face às restantes. Para os primeiros, segundos e terceiros intervalos apenas se destacam as posições iniciais e as restantes não se diferenciam entre si.

Para os restantes intervalos assiste-se a uma certa uniformidade no valor dos *ratings* comerciais.

(4) A hora de transmissão dos programas a que está associado os maiores valores de *ratings* comerciais corresponde às 20H (14,3%), às 22H (9,6%), às 23H (8,8%), às 19H e às 21H (8%). Um segundo pico de valores mas muito menor observa-se às 13H (5%). A tarde e a madrugada são os períodos com menores valores de *ratings* com o valor de 2,7% e 3,3% respectivamente. Seguem-se o pré-prime-time e o princípio tarde com os valores médios de 4,9% e 4,1% respectivamente e com maiores valores o prime time e a noite com os valores de 9,3% e 9,7% respectivamente.

(5) Verificamos que no princípio da tarde e no decorrer da tarde existe uniformidade de *ratings* para todas as localizações. No pré-prime-time existem maiores valores de *ratings* associados às posições intermédias do primeiro espaço comercial em relação às restantes localizações e os restantes intervalos apresentam uniformidade de valores de *ratings*. Para os períodos de prime time e noite destacam-se as duas posições iniciais para todas as localizações. Durante a madrugada existe uma redução nas posições intermédias para os intervalos que antecedem os programas e para os primeiros e alguma uniformidade para as restantes localizações.

(7) As promoções de programas não apresentam maiores valores de *ratings* ($Rat\_Spot\_Médio=5,9\%$) face às restantes

inserções (5,5%). As inserções precedidas por promoções também não apresentam maiores valores de *ratings* com um valor médio de 5,9% e de 5,5% para o conjunto complementar.

(8) Os *ratings* comerciais mais elevados correspondem aos programas de Bolsa e Finanças, Telejornal, Meteorologia, Filmes de Fantasia, Comédias e Histórico e Épicos segundo as tipologias de Grau 3. Atendendo às tipologias de Grau 2 os *ratings* mais elevados estão associados à transmissão de eventos desportivos, informação especializada, séries, telejornal e telenovelas. Portanto, em relação às audiências comerciais estas são superiores uma vez mais para os filmes, programas desportivos e programas de muito curta duração adjacentes aos espaços informativos. As telenovelas, os espaços informativos e as séries ainda que apresentem valores mais reduzidos de *Spot_Ratio%* em termos absolutos não deixam de ser programas a ter em consideração.

Ic. Resultados para o Canal SIC

A SIC também é um canal comercial o que influencia a sua grelha de programas e a estruturação dos espaços comerciais. Recorrendo sobretudo a telenovelas e aos restantes programas que são transmitidos em episódios pretendem a fidelização da audiência nos períodos de maior audiência. A base de dados comercial deste canal está constituída por 17.789 observações. Neste canal reduzem a incidência em inserções antes do início dos programas para não reduzir os fluxos de audiências para os canais concorrentes e apostam nos primeiros espaços comerciais.

Os primeiros intervalos são em menor número, mas existe neles uma maior concentração de inserções e nos espaços comerciais que antecedem os programas verifica-se o inverso, são mais reduzidos.

➤ **Grau de Associação existente entre determinadas variáveis nominais e a variável Dependente *Spot_Ratio%***

(1) A variável *Spot_Ratio%* é maximizada para os espaços localizados antes do início dos programas, assim como para os primeiros e quartos intervalos (0,83). Os segundos e os terceiros intervalos são os mais penalizados (0,7). Os programas mais curtos com apenas um intervalo ou os intervalos finais dos programas apresentam maiores médias e os efeitos de herança favorecem os espaços que antecipam o início dos programas.

(2) A primeira posição (0,95) e a segunda (0,89) são as mais favorecidas e as restantes posições não se diferenciam (0,82).

(3) Cruzando as variáveis localização do espaço comercial e posição verificamos que para os intervalos que antecedem os programas, a primeira e a segunda posição são as que mais retêm a audiência. Verifica-se uma significativa redução nas posições intermédias e as restantes não se diferenciam, ocupando valores intermédios. Para os primeiros, segundos e terceiros intervalos apenas prevalecem as duas posições iniciais. Para o quarto intervalo existe uma tendência decrescente muito acentuada no seu decorrer.

(4) As inserções do final da tarde ou da tarde apresentam uma maior média (0,87), segue-se o final da noite conjuntamente com a noite (0,77) e a madrugada apresenta o menor valor (0,56).

(5) Cruzando os factores período horário, localização e posição foi possível retirar as seguintes conclusões:

No decorrer da tarde os intervalos associados à primeira parte apresentam uma maior estabilidade de valores. A primeira posição destaca-se para os segundos, terceiros e quartos intervalos; a segunda posição apenas para os segundos intervalos e as últimas posições são as que apresentam os menores valores para os quartos intervalos.

No princípio da tarde apenas os intervalos que antecedem os programas apresentam alguma estabilidade no seu desenrolar e para os restantes destacam-se as duas posições iniciais.

Para o pré-prime-time os intervalos que antecedem os programas apresentam uma redução para as posições intermédias e os primeiros intervalos um ligeiro acréscimo para a posição final. Nos segundos e terceiros intervalos destacam-se as duas posições iniciais e as restantes que apresentam uma forte redução, não se diferenciam entre si. Para o quarto espaço comercial verifica-se o mesmo, mas a última posição apresenta um acréscimo.

No prime time os intervalos que antecedem os programas apresentam uma redução nas posições intermédias e os primeiros intervalos apresentam um ligeiro acréscimo nas duas posições iniciais.

Para o período da noite as duas posições iniciais sobressaem significativamente em relação às restantes para os intervalos que antecedem os programas e ligeiramente para os primeiros. Os segundos intervalos apresentam uma maior redução de valores de *Spot_Ratio%* nas posições intermédias.

Para a madrugada e para os espaços que antecedem os programas ou os primeiros intervalos sobressaem significativamente as duas posições iniciais e para os restantes também, mas de forma mais moderada. As restantes posições não se diferenciam.

(6) As inserções que não são promoções (0,81 face a 0,77 de média para os que o são) e as que não são precedidas por elas apresentam maiores valores (0,81 face a 0,78). Apenas para este canal as promoções influenciam este indicador.

Na bibliografia da especialidade alguns autores defendem que as promoções de programas cativam em maior grau as audiências, logo as inserções que as precedem também seriam favorecidas. Para este canal a nível médio e atendendo ao horizonte temporal em análise não se observa esta tendência.

➤ **Grau de Associação existente entre determinadas variáveis nominais e a variável Dependente *Rat_Spot%***

As inserções localizadas entre programas apresentam uma maior média de *Rat_Spot%* com 7% face aos 5% para os intervalos no decorrer dos programas, uma vez que as inserções que correspondem aos segundos, terceiros e quartos intervalos são muito penalizadas a nível comercial (3,5%, de média para *Rat_Spot%*). As inserções com maiores valores de audiências médias correspondem às localizados antes do início dos programas (7% para *Rat_Spot%*) ou no primeiro intervalo (6%), independentemente de este ser o único ou não. Os telespectadores permanecem em frente ao televisor para ver o início dos programas adjacentes, favorecendo os espaços comerciais entre os programas.

(1) A primeira e a segunda posição de um espaço comercial são as que apresentam maior valor médio (6,3% para *Rat_Spot%*). Entre as posições intermédias, antepenúltima, penúltima e última não se observam diferenças significativas (5,7% de média para *Rat_Spot%*). Como se constatou para os restantes canais, em geral as posições iniciais são as que apresentam maiores valores de ratings, uma vez que ao longo do espaço comercial existe uma maior tendência para o zapping. Por vezes, regressam ao programa inicial após ter passado um certo intervalo de tempo favorecendo pontualmente as posições finais. No entanto, para este canal a última e a penúltima posições apresentam apenas um ligeiro acréscimo a nível global face às restantes posições, não considerado

estatisticamente significativo com base na metodologia Anova e para um valor de significância de 5%.

(2) Cruzando os dois factores anteriormente mencionados, verificamos que os espaços comerciais que antecedem os programas apresentam maiores valores de *ratings* para as posições iniciais e para a penúltima e para as restantes localizações apenas se destacam as duas posições iniciais.

(3) As inserções da madrugada são as que minimizam os valores de *Rat_Spot%*. Segue-se a Tarde (3,8%), o Princípio da Tarde (5,6%), o pré-prime-time (5,9%), a Noite (7,8%) e com maiores valores o Prime Time (9,3%).

(3) Cruzando os factores período horário, a localização do espaço comercial e a posição da inserção constatamos que no início da tarde, os espaços comerciais que antecedem os programas apresentam uma redução nas posições intermédias. Para as restantes localizações destacam-se apenas as posições iniciais. No decorrer da tarde existe uniformidade de valores de audiências comerciais no desenrolar dos espaços comerciais, para todas as localizações.

No pré-prime-time para os espaços comerciais que antecedem os programas existe uma forte redução nas posições intermédias e as restantes não se diferenciam entre si. Para os primeiros espaços comerciais verifica-se também uma redução nas posições intermédias, mas mais ligeira e para os restantes espaços comerciais destacam-se as posições iniciais.

No período do prime time os espaços comerciais que antecedem os programas e os primeiros intervalos apresentam uma ligeira redução de *Rat_Spot%* nas posições intermédias e na antepenúltima.

Para o período da noite os intervalos que antecedem os programas têm maiores valores de *Rat_Spot%* face aos restantes intervalos e destacam-se as duas posições iniciais para todos eles.

(5) Os inserções que não correspondem a promoções de programas com um valor médio de 5,7% ou de 4,8% para o complementar e os que não são precedidos por elas têm uma maior média no valor de 5,6% ou de 4,7% para o complementar. Para a SIC não se constata que as promoções de programas sejam mais apelativas ao público face às restantes inserções comerciais.

(6) Em relação às tipologias programáticas de Grau 3 podemos ordená-las da seguinte forma: As tipologias que apresentam uma maior média de *Rat_Spot%* são os *sketch humorísticos*, os telejornais, os *talk shows* e a entrega de prémios com valores no intervalo [5.3%; 9.2%]. Seguem-se a maior parte dos filmes na seguinte ordem decrescente de ratings: western, policial, beleza, romance, aventura, família, comédia, acção, natureza, guerra e drama com valores no intervalo [2.8%; 5.1%]. Com menores valores de ratings destacam-se por ordem decrescente as reportagens, informação cultural, stand up comedy, filmes de suspense e ficção científica, séries e música (valores no intervalo [0.8%; 2.5%]).

Os programas que mais retêm as audiências são os de curta duração após os espaços informativos da noite, o jornal da tarde de Segunda a Sexta, as telenovelas e as séries. Os espaços informativos e os seus programas adjacentes além de apresentarem maiores índices de fidelidade, também são os que apresentam maiores valores de índices de audiências comerciais.

➤ **Análises a Valores de Espaços comerciais**

Concluídas as análises de avaliação da associação de determinados factores com os valores médios de *Rat_Spot%* ou de *Spot_Ratio%* prosseguimos com as análises realizadas aos espaços comerciais por aplicação do algoritmo de *clustering* VQ que corresponde a uma Rede Neuronal de *clustering* com um funcionamento próximo ao K Média e por aplicação do algoritmo de *clustering* DIV simbólico. Pretende-se identificar quais são os espaços comerciais com maiores valores médios de ratings, mas também com uma maior estabilidade de valores.

Com base nas análises de *Data Mining*, a melhor solução corresponde a cinco *clusters*.

(1) O *cluster* que maximiza os valores médios, mas que também apresenta apenas uma unidade percentual de variabilidade de valores, numa elevada percentagem corresponde aos primeiros e segundos intervalos da principal telenovela de prime time. A popularidade das telenovelas da noite incute regularidade nas correspondentes audiências comerciais. Quanto aos espaços comerciais antes do início dos programas estes estão associados a vários

programas. Atendendo aos períodos horários, correspondem a espaços comerciais entre os programas de prime time e em menor percentagem da noite, aos primeiros intervalos destes dois períodos com percentagens semelhantes e aos segundos intervalos dos programas da noite. Para os intervalos localizados antes do início dos programas ou os primeiros correspondem em mais de 75% aos dias laborais e quanto aos segundos metade corresponde a espaços de fim-de-semana.

(2) O *cluster* que se segue com os maiores valores médios, mas apresentando uma grande variabilidade de valores, corresponde aos intervalos essencialmente do período da noite independentemente da sua localização. Em menor percentagem também constam da madrugada para os segundos intervalos e de prime time para as restantes localizações. Quanto aos programas a que correspondem, apenas para os segundos intervalos existe uma significativa percentagem associada à principal telenovela da noite, sendo esta muito inferior para o primeiros intervalos. Correspondem aos dias laborais em 75% das observações.

(3) O *cluster* que apresenta uma média ligeiramente inferior, mas com uma reduzida variabilidade de valores, atendendo aos espaços comerciais entre programas ou para os primeiros intervalos, corresponde em metade das observações a intervalos do prime time, aos segundos intervalos da noite e aos terceiros da tarde. Em relação aos programas, as duas primeiras localizações correspondem em mais de 15% a intervalos do espaço informativo da noite.

(4) O *cluster* que apresenta não só os menores valores, mas também uma maior estabilidade, corresponde a espaços dos períodos da tarde e da madrugada, incrementando a percentagem deste último período com a localização do espaço comercial.

Realizaram-se idênticas análises com os dados simbólicos, sendo os espaços comerciais agregados em função do programa, localização do espaço comercial e período em que se diferencia a semana do fim-de-semana. Por aplicação do algoritmo de *clustering* simbólico divisivo e utilizando como variável de classificação apenas o valor médio de *Rat_Spot%* verificamos que para uma solução a seis *clusters* a percentagem de variância explicada é de 90%.

Ao *cluster* com maiores valores corresponde todos os espaços comerciais associados à principal telenovela da noite de semana e de fim de semana; aos espaços comerciais de programas de muito curta duração; ao fim-de-semana ao primeiro espaço comercial do Primeiro Jornal e aos espaços comerciais dos programas sucessores ao espaço informativo da noite.

Segue-se o *cluster* que inclui os espaços comerciais do espaço informativo, da segunda telenovela da noite, trofeus do desporto e ao fim-de-semana aos espaços comerciais que antecedem os programas destinados à transmissão de filmes ao início noite ou após o espaço informativo da tarde, o programa de humor sucessores do espaço informativo e a segunda telenovela da noite.

Portanto, os maiores valores estão associados aos programas de prime time e do início da noite.

II. Análises de Regressão Linear Múltipla

Neste subcapítulo constam os resultados das análises lineares por aplicação da Regressão Linear Múltipla ou dos equivalentes algoritmos simbólicos. As análises permitiram avaliar o grau de linearidade das relações entre as variáveis e determinar se a previsão dos *ratings* comerciais pode ser realizada por esta via.

Como se verificou nas análises exploratórias os canais diferenciam-se quanto às variáveis mais relevantes, o que pode ser atribuído à composição sociodemográfica da sua audiência, ao planeamento publicitário, à estruturação da grelha de programas, à imagem do canal, entre muitos outros factores. Consequentemente, as análises realizam-se para cada um dos canais, ou diferenciando canais públicos (RTP1 e Canal 2) e canais privados (SIC e TVI). Também foram realizadas análises para a totalidade da emissão diária, assim como, para alguns períodos horários mais restritos.

IIa. Resultados para o Canal SIC

O gráfico F.7.5.1 ilustra os valores médios e os valores de variabilidade dos *ratings* em função das localizações dos espaços comerciais e das faixas horárias.

O maior valor de 7% é do período da noite, seguindo-se o horário nobre. Considerando os canais estatais o maior valor de variabilidade de *ratings* corresponde a 18% no decorrer do horário nobre devido essencialmente à transmissão dos eventos desportivos. Se excluirmos as observações atípicas a variabilidade não excede os 6%. Logo, as análises de previsão de *ratings* talvez possam produzir resultados razoavelmente precisos, para alguns dos canais e faixas horárias.

Na *Tabela F.7.5.1* constam os valores dos coeficientes de determinação sobre uma amostra de teste que corresponde a 30% das observações para os modelos de Regressão Linear Múltipla não Simbólica. Para alguns períodos e considerando a variável critério *Rat_Spot%* os valores dos coeficientes de determinação podem ser razoáveis. Inicialmente apresentam-se os resultados das análises de correlação, seguindo-se a descrição das análises de Regressão Linear Múltipla e de RLM simbólica considerando como variável dependente *Rat_Spot%* e *Spot_Ratio%*.

Variância Explicada	Rat_Spot (%)	Spot_Ratio (%)
Canais Públicos		
Global	0,82	0,25
Prime Time + Noite	0,76	0,48
Pre Prime Time	0,52	0,35
Tarde	0,76	0,53
Canais Privados		
Global	0,87	0,22
Prime Time + Noite	0,85	0,43
Pré Prime Time	0,54	0,31
Tarde	0,8	0,52

Tabela F.7.5.1 – Coeficientes de determinação dos modelos de Regressão Linear Múltipla

Variáveis Explicativas	Rat_Spot (%)	Spot_Ratio (%)
Sat_Ldin%	-0,48	0,10
Num_Partes_Total	-0,46	-0,20
Num_Promos	-0,38	-0,16
Dur_Ldin	-0,35	0,10
Sat_Prog%	-0,31	0,02
Dur_Prog (T)	-0,28	-0,22
Dur_Início_Parte	-0,19	-0,18
Ratio_Início_Parte	-0,10	-0,13
Dur_Spot	-0,01	-0,03
Dur_Ac	-0,017	-0,13
Ratio_Dur_Ac	-0,07	-0,11
Hora_Spot	0,64	0,28
Rat_Prog (P)	0,87	-0,21
Rat_Prog (T)	0,75	-0,04
Tvr_Prog (T)	0,74	-0,16
Rat_Ldin%	0,57	-0,21
Fid_Prog (T)%	0,56	0,20
Rat_Ldout%	0,48	0,09
Hora_Inicio_Prog (T)	0,46	-0,03
Shr_Prog% (T)	0,21	0,06
Fid_Prog% (P)	0,36	0,22

Tabela F.7.5.2 – Coeficientes de correlação linear de pearson para o canal SIC

A *Tabela F.7.5.2* e a *Tabela F.7.5.3* resumem alguns dos coeficientes de correlação linear entre as duas variáveis critério e as restantes variáveis explicativas e eventualmente de previsão.

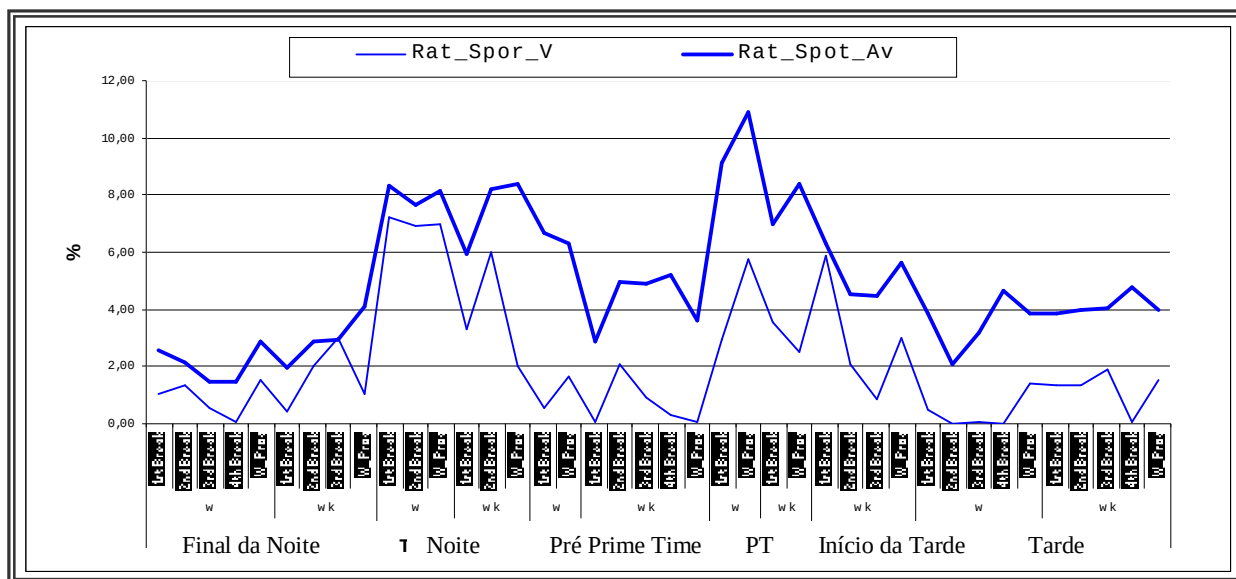


Gráfico F.7.5.1 – Evolução de ratings comerciais para o canal SIC

➤ Coeficientes de Correlação linear para *Rat_Spot%*

Com base nos valores de correlação linear de pearson verificamos que a maioria das variáveis apresenta valores reduzidos, o que indicia em geral a falta de linearidade das relações existentes, mas, não, uma ausência de relação.

A saturação publicitária de um programa ou do seu antecessor e uma elevada duração, quer do programa anterior, quer do programa associado à inserção penalizam os valores dos ratings comerciais, mas prevalece a influência do programa anterior.

Os espaços comerciais com um maior número de promoções, também correspondem aos mais penalizados. Por outro lado, o *rating* do programa global e parcial condicionam positivamente esta variável, assim como também, a audiência média para o total da televisão. O *rating* do programa anterior e sucessor, um horário mais tardio, uma maior fidelidade ao programa e uma maior duração da parte do programa associada ao espaço comercial são também variáveis a considerar com uma influência positiva nos valores dos ratings comerciais.

Por outro lado, uma maior duração desde o início do programa até à parte associada ao espaço comercial penaliza esta variável, a nível absoluto e relativo (*Dur_Início_Parte* e *Ratio_Início_parte*).

Variáveis Explicativas	Rat_Spot (%)	Spot_Ratio (%)
Beta_Prog (T)	0,21	0,07
Beta_Prog (P)	0,12	0,01
Ratio_Ldin%	0,14	0,20
Dur_Prog (P)	0,10	-0,24
Dur_BS	0,10	-0,26
Ratio_Dur_BS	0,10	0,03
Ratio_Ldout	0,02	-0,29
Num_Inserções	0,02	-0,14
Ratio_Dur_Fim	0,04	0,07
Dur_Espaço comercial	0,056	-0,11
Dur_Fim	0,07	0,03
Avr		0,29
Bsh_Prog (T)	-0,30	-0,22
Tvr_Prog (P)	0,78	-0,19
Pos_Spot	-0,03	-0,12
Ratio_Dur_Ac	-0,05	-0,10
Hora Início_Prog (P)	0,57	-0,04
Atv		0,23
Dur_Spot	-0,01	-0,03

Tabela F.7.5.3 – Coeficientes de correlação linear de pearson para o canal SIC

As restantes variáveis apresentam coeficientes muito reduzidos o que apenas nos permite concluir a respeito de ausência de linearidade das relações entre as variáveis, mas, não, a respeito de uma total inexistência de relação.

Por exemplo, uma maior duração desde o início do espaço comercial até à inserção a nível absoluto e relativo e a duração da inserção correspondem a variáveis que apresentam valores irrisórios, mas negativos. Os maiores valores de correlação estão associadas ao *rating* programático parcial (0,87) e ao *rating* total (0,74), assim como às taxas médias de audiências para o total da TV parciais e totais (0,78 e 0,74 respectivamente). Constatamos uma maior associação com os indicadores parciais uma vez que abrangem um período mais restrito.

➤ Os Modelos de Regressão Linear para Rat_Spot%

(1) Resultados Considerando a Totalidade da Amostra

Se incluirmos nas análises como variável independente apenas os valores dos *ratings* programáticos parciais, o R^2 apresenta o valor de 0,79 (beta=0,89). Portanto, os programas com maiores níveis de exposição média, também, apresentam maiores audiências comerciais, existindo uma razoável associação linear entre estas duas variáveis. Se restringirmos as análises a determinados períodos horários o grau de linearidade existente poderá variar.

O modelo com um maior coeficiente de determinação inclui 20 variáveis para conseguir um acréscimo de apenas 6% na percentagem de variância explicada e com um R^2 no valor de 0,85.

Tentando identificar um modelo alternativo mais parcimonioso, por exemplo, se incluirmos como variáveis explicativas, para além do *rating* programático parcial ($Beta=0,82$), o número de promoções ($beta=-0,14$) e a duração do programa (P) ($Beta=-0,12$) obtemos um R^2 ajustado de 0,80. Não se justifica incluir mais duas variáveis para um acréscimo de apenas 0,1. Obviamente, que o número de promoções, o que acarreta uma maior durabilidade do espaço comercial, como também a duração programática penalizam as audiências.

Incluindo apenas o valor de *Rat_Prog% (P)* e *Tempo_Spot* que representa a hora da inserção transformada num número decimal elevamos o R^2 para 0,82 (+3%) com betas no valor de 0,80 e 0,21 respectivamente.

Se em vez da hora da inserção a substituíssemos por *Time_Prog (P)* que corresponde à hora de início do programa transformada num número decimal, o valor de R^2 seria menor e com um valor de 0,80 uma vez que as inserções pertencentes a um mesmo espaço comercial podem apresentar consideráveis variabilidades nos seus valores em função da posição que ocupam.

Se incluirmos *Rat_Prog% (P)*, *Tempo_Spot* e *Dur_Ac* para ter em consideração a posição da inserção o valor de R^2 permanece inalterável com esta última variável a apresentar um beta no valor de -0,07 e o mesmo se verifica para *Pos_Spot* ou *Dur_Espaço_Comercial*.

As posições intermédias num espaço comercial são as mais penalizadas e as primeiras ou as últimas são mais favorecidas, logo a relação com *Dur_Ac* não é linear. Parece preferível associar *Rat_Prog%* e *Tempo_Spot* à variável *Dur_BS* ($beta=-0,13$), mas o acréscimo é de apenas 1% ($R^2=0,83$) e se em vez desta variável associarmos *Tvr_Prog% (P)* o beta desta variável não é significativo. Se incluirmos *Tempo_Spot* e *Tvr_Prog% (P)* o R^2 é de 0,67 com betas no valor de 0,30 e 0,61 respectivamente. Portanto, com *Tvr_Prog%* é menor a associação do ponto de vista linear.

Experimentando outros modelos não é possível maximizar o coeficiente de determinação para além dos 0,83. Interessa, portanto, restringir as análises a determinados períodos horários por forma a verificar a validade destes resultados, sobretudo para os períodos de maiores audiências. O modelo mais parcimonioso será o que inclui apenas *Rat_Prog% (P)* considerando a totalidade dos dados.

(2) Resultados Considerando a Amostra Referente ao Prime Time e à Noite

Verificamos que existe um decréscimo no grau de associação linear existente entre *Rat_Spot%* e *Rat_Prog% (P)*, sendo o R^2 de 0,59 e o $beta=0,76$. Nos períodos de maiores audiências a relação entre os *ratings* programáticos e os valores comerciais afasta-se da linearidade. O acréscimo verificado nos *ratings* não é acompanhado por subidas proporcionais nos respectivos espaços comerciais.

Isto denota o carácter mais activo do público aproveitando os espaços comerciais para mudar de canal ou realizar outras actividades. Consequentemente, a posição da inserção no espaço comercial é determinante e condiciona o seu grau de exposição.

Se incluirmos também *Dur_BS* elevamos o R^2 para 0,66 (+7%) e com um valor de beta de -0,27. Como consequência da actividade do público quanto mais tempo decorrer desde o início do programa até à transmissão da inserção mais penalizados serão os mesmos.

Incluindo também *Tempo_Time* (*P*) o R^2 é de 0,70 (beta=-0,22), ou seja, quanto mais avançamos no decorrer da noite menores serão as audiências programáticas e as comerciais. Que o sejam em valor nada significa que também o sejam a nível relativo. Para esta ultima avaliação será necessário recorrer a *Spot_Ratio%*.

Incluindo *Dur_Espaço_Comercial*, o R^2 atinge o valor de 0,73 (beta=-0,17) e, logo, as durações dos espaços comerciais penalizam as audiências comerciais. Por último se incluirmos *Tempo_Spot* e *Tvr_Prog%* (*P*) ambas com coeficientes negativos o R^2 atinge os 0,75.

Portanto é necessário incluir cinco variáveis para um acréscimo de 0,16 no coeficiente de determinação. Este modelo inclui *Rat_Prog%* (*P*), *Dur_BS*, *Tempo_Prog* (*P*), *Dur_Espaço_Comercial*, *Tempo_Spot* e *Tvr_Prog%* (*P*).

Para conseguirmos alcançar 75% de variância explicada é necessário recorrer aos *ratings* programáticos e ao seu cálculo poderá estar associado grandes margens de erro.

Se a excluíssemos das análises, incluindo as restantes variáveis que foram enumeradas o coeficiente de determinação seria de apenas 0,30, sendo a variável *Tvr_Prog%* (*P*) a que apresenta um maior valor de beta e positivo com o valor de 0,48. Portanto, do ponto de vista linear é grande a interdependência entre as audiências comerciais e as programáticas.

(3) Resultados Considerando a Amostra Referente ao Pré-Prime-Time

O grau de associação linear entre *Rat_Spot%* e *Rat_Prog%* é de 0,25, bastante inferior ao valor que se verifica para os períodos de maiores audiências. Agregando *Tempo_Spot* sobe para 0,41 e com *Tempo_Prog* para 0,26. Se esta variável fosse a única o R^2 seria de apenas 0,09. Não existe uma associação linear entre a hora da inserção publicitária e os *ratings* comerciais. Tínhamos verificado que existe uma súbita exponencial ou polinomial nos valores dos *ratings* programáticos para este período, logo dificilmente existiria linearidade a nível comercial. Com as variáveis *Rat_Prog%* (*P*), *Tempo_End* (*P*), *Dur_BS*, *Pos_Spot*, *Dur_Ac* conseguimos um R^2 de 0,52. Apenas os *ratings* programáticos e a duração acumulada estão positivamente correlacionadas com *Rat_Spot%*. As variáveis *Dur_BS* e *Pos_Spot* penalizam os *ratings* uma vez que as posições intermédias apresentam fortes reduções nos seus valores. Inclusive a hora de conclusão do programa penaliza as inserções, ou seja, o acréscimo verificado nos *ratings* programáticos não será necessariamente acompanhado por uma idêntica repercussão a nível comercial.

Portanto, para este período as relações existentes afastam-se da linearidade.

(4) Resultados considerando a amostra referente ao decorrer da tarde

No decorrer da tarde verificam-se os maiores valores de associação linear entre *Rat_Spot%* e *Rat_Prog%* (*P*) sendo o R^2 de 0,63 e o $\beta = 0,76$ e que corresponde a um valor superior ao do horário nobre e da noite. Se incluirmos *Tempo_Spot* o valor do coeficiente de determinação ascende a 0,65 e adicionando *Tempo_Prog* ascende a 0,67. O valor do β para a hora da inserção é positivo, mas com a hora de início do programa é negativo. No decorrer da tarde existe uma redução nos ratings programáticos e, consequentemente também deveria ser esta a tendência a nível comercial, logo o valor positivo de *Tempo_Spot* contradiz as tendências identificadas.

Com as variáveis *Rat_Prog%* e *Tvr_Prog%* ($\beta = -0,39$) o R^2 é de 0,70. Logo, é preferível incluir as taxas médias de audiências para o total da TV em vez da variável hora. No entanto maiores níveis de *Tvr_Prog%* (*P*) penalizam as audiências, mas esta variável associa-se positivamente com *Rat_Prog%*. Em princípio deveria verificar-se o inverso, ou seja, subindo os valores programáticos, os espaços comerciais, também deveriam ser beneficiados.

Considerando apenas *Tvr_Prog%* (*T*) o seu coeficiente é positivo e se incluirmos também *Rat_Prog%* o R^2 é de 0,66 e inferior ao de considerar *Tvr_Prog%* (*P*).

Considerando combinações das restantes variáveis os modelos não atingem os 70% de variância explicada.

O melhor modelo corresponde consequentemente a *Rat_Prog%* e *Tvr_Prog%* ($\beta = -0,39$) com um R^2 no valor de 0,70. Estes resultados para o decorrer da tarde resultam de considerar a totalidade do período semanal.

Diferenciando os dias laborais do fim-de-semana, verificamos que durante a semana o valor de R^2 com *Rat_Prog%* é de 0,68 e, logo superior ao valor global. Considerando as variáveis *Rat_Prog%* e *Tempo_Spot* e ambas com coeficientes positivos o R^2 é de 0,80. Para atingirmos os 85% de variância explicada é necessário incluir também *Dur_Espaço_comercial* e *Dur_Ac* com coeficientes positivos e *Dur_BS* e *Tempo_Time* com coeficientes negativos. Nos dias laborais existe uma acentuada descida nos valores dos ratings programáticos no decorrer da tarde. Logo, a hora de início deve penalizar as inserções, assim como uma maior duração desde o início do programa até à inserção.

As posições iniciais ou finais às que se associa um maior valor de *Dur_Ac*, têm maiores valores de ratings e em relação à duração do espaço comercial o seu coeficiente é de apenas 0,085. Apesar da significância estatística deste resultado o seu valor é irrisório e pouco contribui para a variância explicada pelo modelo. O modelo mais parcimonioso corresponde, portanto a *Rat_Prog%* e *Tempo_Spot*.

Considerando apenas o fim-de-semana verificamos que o valor do R^2 com *Rat_Prog%* é de 0,58 e inferior ao global e ao calor do modelos dos dias laborais. Se incluirmos *Tempo_Spot* o resultado permanece inalterado, ao contrário do que se verificou para os dias laborais.

Associando *Rat_Prog%* a *Tempo_Prog* existe um acréscimo de 3% ($R^2=0,61$), apresentando esta última variável um coeficiente negativo. O melhor resultado e mais parcimonioso incluem *Rat_Prog%* e *Tvr_Prog%* (*T*) ($R^2=0,70$) com os *ratings* programáticos a favorecer *Rat_Spot%* e com as taxas médias de audiências para o total da TV a penaliza-los. Ao fim-de-semana no decorrer da tarde é menor a associação linear entre as variáveis.

➤ Análises de Correlação para *Spot_Ratio%*

Atendendo à variável *Spot_Ratio%* (Tabela F.7.5.4) verificamos que uma maior duração programática global ou parcial favorece os comportamentos de *zapping*, assim como os espaços comerciais mais afastados do início dos programas a nível absoluto e relativo ou inseridos em programas com um maior número total de partes. As restantes variáveis a considerar são o *Rat_Ldin%*, o *Ratio_Ldout%*, o *rating* programático parcial e global, o valor de *Tvr_Prog%* total e parcial, as quais também influenciam negativamente esta variável. Os períodos com maiores valores das audiências para o total da TV e de valores programáticos apresentam uma menor proporção de *ratings* comerciais face aos programáticos, o que poderá ser atribuído, quer a uma maior intensidade no *zapping*, quer também talvez a um possível comportamento da audiência semelhante aos restantes períodos, mas penalizado pela existência de maiores valores de *ratings* programáticos. O número de promoções por espaço comercial, a duração acumulada até à emissão da inserção a nível absoluto e relativo, o número de inserções por espaço comercial, a posição da inserção no espaço comercial e a duração do espaço

comercial também estão negativamente correlacionadas com *Spot_Ratio%*.

Variáveis	Spot_Ratio (%)	Rat_Spot (%)
Ratio_Inicio_Parte	0,02	-0,17
Num_Parties_Total	0,03	-0,30
Num_Parte_Prog	0,03	-0,30
Ratio_BS	0,03	-0,04
Fid_Prog (T)	0,03	0,68
Num_Inserções	0,04	-0,19
Dur_BS	0,05	-0,18
Rat_Spot%	0,05	
Fid_Prog (P)	0,05	0,44
Dur_Inicio_Parte	0,08	-0,20
Sat_Prog%	0,01	-0,21
Dur_Fim	0,10	0,002
Ratio_Dur_Fim	0,11	0,05
Num_Promos	0,12	0,12
Rat_Ldout%	0,26	0,46
Hora_Inicio_Prog (G)	0,30	0,54
Hora_Inicio_Prog (P)	0,31	0,52
Ratio_Ldin%	0,36	0,14
Hora_Spot	0,36	0,59

Tabela F.7.5.4 – Coeficientes de correlação para o canal SIC e atendendo à variável *Spot_Ratio%*

Todas as variáveis associadas a uma maior duração do espaço comercial e do programa, o desempenho do programa ou dos adjacentes e os períodos com maiores níveis de *Tvr%* penalizam os *ratings* comerciais quando

medidos relativamente aos índices programáticos. As variáveis que apresentam coeficientes positivos ainda que de reduzida magnitude correspondem ao *Sat_Ldin%* e a *Dur_Ldin*, à fidelidade ao programa e à hora da inserção. Uma maior fidelidade ao programa e a hora da inserção favorecem ambas as variáveis. Com *Sat_Ldin%* e *Dur_Ldin* verificam-se efeitos contraditórios, ou seja, tendem a favorecer *Spot_Ratio%*, mas penalizam *Rat_Spot%*. Os *ratings* programáticos parciais e totais, assim como dos programas adjacentes, o *Tvr_Prog%* (T/P), a duração do programa (P) e a hora de início do programa favorecem *Rat_Spot%*, mas penalizam *Spot_Ratio%*. Os maiores valores de correlação correspondem a *Hora_Spot* (0,28) e *Ratio_Ldout* %(-0,29) e *Dur_BS* (-0,26).

➤ Os Modelos de Regressão Linear para *Spot_Ratio%*

Atendendo à variável *Spot_Ratio%* o modelo de Regressão Linear para a totalidade da emissão diária com um maior coeficiente de determinação, apresenta um R^2 de apenas 0,23 e inclui inúmeras variáveis. Incluí as variáveis *Dur_Prog* (P) ($\beta=-,03$), *Rat_Ldin%* ($\beta=,081$), *Num_Promos* ($\beta=-,23$), *Dur_Prog* (T) ($\beta=-,26$), *Rat_Prog%* (P) ($\beta=-,26$), *Rat_Ldout%* ($\beta=,209$), *Tvr_Prog%* (T) ($\beta=-,27$), *Hora_Início_Prog* ($\beta=,112$), *Dur_Ac* ($\beta=-,10$), *Dur_Ldin* ($\beta=,158$), *Sat_Ldin%* ($\beta=-,13$), *Sat_Prog%* ($\beta=,107$), *Num_Inserções* ($\beta=,045$), *Dur_Spot* ($\beta=-,03$). Com as variáveis *Tempo_Spot*, *Rat_Prog%*, *Dur_Ac*, *Dur_Prog* (P), *Tempo_End* (T) o valor do R^2 é de 0,22 e apenas com *Tempo_Spot* apresenta um valor positivo.

Portanto, horas mais tardias apresentam maiores valores de *Spot_Ratio%* ainda que esta associação não seja linear. O *rating* de um programa penaliza *Spot_Ratio%* uma vez que está em denominador quando se calcula esta variável. Uma maior duração acumulada pode implicar menores valores de *ratings* comerciais, assim como maiores durações programáticas parciais uma vez que o público pode utilizar os espaços comerciais para realizar outras actividades. O valor de *Tempo_End* também as penaliza, mas apresenta um coeficiente irrisório (-0,09). Considerando a globalidade dos dados não existe linearidade entre as variáveis.

Restringindo as análises aos períodos de prime time e noite e se considerarmos as variáveis enumeradas no último parágrafo, com apenas *Tempo_End* que apresenta um coeficiente negativo, o valor do R^2 é de 0,22. Aumentou significativamente o grau de linearidade restringindo as análises a um período horário menos abrangente. No decorrer da noite as audiências programáticas diminuem, mas as audiências comerciais reduzem-se mais que proporcionalmente. Com as variáveis referidas no parágrafo anterior e em que todas apresentam coeficientes negativos o R^2 é de 0,43 e de 0,22 para a totalidade da emissão diária. Portanto, as durações parciais, os tempos das inserções e de conclusão do programa, o *rating* do programa e a duração acumulada penalizam *Spot_Ratio%*. Para os restantes períodos horários verifica-se a mesma ausência de linearidade entre as variáveis.

Com base nas análises realizadas ao canal SIC podemos concluir que os modelos lineares parecem ser uma opção a considerar para o desenvolvimento de modelos explicativos e eventualmente de previsão dos *ratings* comerciais.

Contudo, quando se pretende analisar a variável *Spot_Ratio%* deve-se recorrer a outras opções. Os factores temporais influenciam o grau de associação linear entre as variáveis e a ordenação das mesmas em função da sua capacidade explicativa e de previsão.

Iib. Resultados para o Canal TVI

Na *Tabela F.7.5.5* consta o resumo dos coeficientes de correlação linear de pearson entre as variáveis *Spot_Ratio%* e *Rat_Spot%* e as restantes variáveis explicativas e eventualmente de previsão.

➤ Coeficientes de Correlação para *Spot_Ratio%*

Para a variável *Spot_Ratio%* as variáveis que se associam proporcionalmente com a mesma são a duração desde a inserção até à conclusão do espaço comercial a nível absoluto e relativo, o que corresponde às posições iniciais; a posição da inserção, ou seja, as posições mais próximas do final do espaço comercial e os programas com maiores valores de *Rat_Ldout%* e *Ratio_Ldin%* e associados a horas mais tardias.

Com uma influência inversamente proporcional em *Spot_Ratio%* constam *Ratio_Ldout*, *Rat_Ldin%*, *Beta_Prog (P/T)*, *Rat_Prog (P)*, *Tvr_Prog% (P)*, *Dur_Prog (P)*, *Tempo_Início (P)* e *Ratio_Dur_Ac*. Corresponde a programas com maiores valores de *ratings* programáticos ou dos valores adjacentes, inseridos em períodos horários com maiores audiências para o total da TV e com maiores durações parciais.

A duração desde o início do espaço comercial até à inserção a nível relativo, também, penaliza este indicador.

Variáveis	Spot_Ratio (%)	Rat_Spot (%)
Ratio_Ldout%	-0,50	0,23
Rat_Ldin%	-0,48	0,74
Beta_Prog (P/T)	-0,32/-0,23	0,32/0,48
Rat_Prog% (P)	-0,26	0,92
Tvr_Prog% (P)	-0,17	0,85
Dur_Prog (P)	-0,13	-0,01
Ratio_Dur_Ac	-0,11	-0,07
Tempo_Inicio (P)	-0,10	0,46
Dur_Ac	-0,09	-0,10
Tempo_Inicio (T)	-0,08	0,50
Rat_Prog% (T)	-0,08	0,95
Bsh_Prog (P)	-0,06	0,01
Pos_Spot	-0,06	-0,11
Sat_Ldin%	-0,05	-0,08
Dur_Spot	-0,05	0,02
Tvr_Prog% (T)	-0,04	0,83
Dur_Ldin	-0,02	-0,20
Dur_Espaço comercial	0,009	-0,22
Bsh_Prog (T)	0,01	-0,25
Dur_Prog (T)	0,01	-0,25

Tabela F.7.5.5 – Coeficientes de Correlação para o canal TVI

A única variável de planeamento comercial com um maior coeficiente corresponde a uma maior duração desde o início do espaço comercial até à inserção, o que corresponderá às posições intermédias, onde se verificam as maiores reduções nos valores dos índices comerciais.

A maioria das variáveis apresenta coeficientes reduzidos, podendo-se concluir que as variáveis não se inter-relacionam de forma linear. No entanto, o sinal dos coeficientes corresponde ao esperado. Com coeficientes negativos constam *Dur_Ac*, *Rat_Prog%* (*T*), *Bsh_Prog* (*P/T*), *Pos_Spot*, *Sat_Ldin%*, *Dur_Spot*, *Tvr_Prog%* (*T*), *Dur_Ldin*, *Dur_Espaço_Comercial*, *Dur_Prog* (*T*), *Ratio_Início_Parte*, *Num_Partes_Total*; *Num_Parte*, *Ratio_BS* e *Pos_Spot*. Com coeficientes positivos constam as variáveis *Rat_Spot%*, *Sat_Prog%*, *Dur_Início_Spot_Parte* e *Fid_Prog%*.

➤ Modelos de Regressão Linear para *Spot_Ratio%*

Aplicando um modelo de Regressão Linear à variável *Spot_Ratio%* com apenas as variáveis métricas verifica-se a não linearidade das relações. O melhor modelo apresenta um R^2 ajustado de apenas 0,18. O maior valor de Beta está associado a *Tvr_Prog%* (*P*) (-0,52), *Dur_Ac* (-0,23), *Rat_Prog%* (*T*) (0,21) e *Sat_Prog%* (-0,16). As restantes variáveis apresentam coeficientes muito reduzidos e também negativos. Os valores de *Tvr_Prog%* parciais penalizam este indicador, mas os *ratings* globais favorecem *Spot_Ratio%*. Compreende-se facilmente que *Tvr_Prog%* (*P*) penalize o indicador *Spot_Ratio%* uma vez que está positivamente correlacionado com *Rat_Prog%* (*P*). Quanto ao sinal positivo associado a *Rat_Prog%* (*T*) é preferível diferenciar os vários períodos e verificar a validade deste resultado. Em princípio maiores valores de *ratings* globais deveriam estar também associados a maiores valores parciais e, consequentemente penalizar a variável *Spot_Ratio%*.

No entanto, como se verificou anteriormente nem sempre existe estabilidade na evolução dos *ratings* ao longo da duração global de um programa. Verificamos que existem algumas partes com níveis inferiores ao valor global.

Se restringirmos as análises aos períodos de prime time e da noite e incluirmos apenas a variável *Rat_Prog%* (*P*) obtemos um R^2 no valor de 0,28, tendo esta variável um coeficiente negativo (-0,53). Esta variável penaliza *Spot_Ratio%* uma vez que está em denominador e os acréscimos nas audiências programáticas não são acompanhadas por súbidas proporcionais nos valores dos *ratings* comerciais. A segunda variável mais correlacionada e incluída no modelo é *Rat_Prog%* (*T*) ao que equivale um valor de R^2 de 0,44, mas com um coeficiente positivo. Para os *ratings* globais uma vez que abrangem um maior horizonte temporal, quanto maior for o seu valor também será *Spot_Ratio%*.

Incluindo *Rat_Ldin%* (*T*) existe um acréscimo de 6% em R^2 e com um beta positivo. Segue-se a variável *Dur_Ac* que penaliza esta variável e com um $R^2 = 0,54$, *Tvr_Prog%* (*T*) com beta negativo e um $R^2 = 0,56$ e *Rat_Ldout%* com um valor de beta positivo e um $R^2 = 0,63$. O acréscimo decorrente da introdução de mais variáveis é marginal. Portanto, existe uma maior linearidade entre os *ratings* programáticos adjacentes ou não e parciais ou globais e a variável *Spot_Ratio%*. O sinal dos betas para os *ratings* globais é positivo e negativo para os valores parciais. Para *Tvr_Prog%* verifica-se o inverso, mas os coeficientes de correlação são sempre negativos.

Após a análise da variável *Rat_Spot%* será mais fácil justificar estes resultados.

Para estes períodos de maiores audiências se restringirmos as análises ao fim-de-semana uma vez mais a variável mais correlacionada com *Spot_Ratio%* é *Rat_Prog% (P)* (R^2 de 0,25) com um valor de beta negativo, segue-se *Rat_Prog% (T)* (R^2 de 0,55) (+), *Rat_Ldin%* com um valor de beta positivo (R^2 de 0,60), *Dur_Ac* com um valor de beta negativo (R^2 de 0,64), *Tvr_Prog% (T)* com um valor de beta negativo (R^2 de 0,67), *Rat_Ldout%* com um valor de beta positivo (R^2 de 0,70), *Sat_Prog%* com um valor de beta positivo (R^2 de 0,71). As variáveis mais correlacionadas correspondem às identificadas anteriormente para a totalidade da amostra e mantém-se o sinal dos betas. Inclui as variáveis associadas a ratings programáticos ou a *Tvr_Prog%* e como variáveis de planeamento apenas inclui a duração acumulada até à inserção. A variável responsável pelo maior acréscimo no coeficiente de determinação corresponde a *Rat_Prog% (T)*.

Realizando idênticas análises para os dias laborais a primeira variável incluída é *Rat_Prog% (P)* (R^2 de 0,32), *Rat_Prog% (T)* (R^2 de 0,48), *Rat_Ldin%* (R^2 de 0,53), *Dur_Ac* (R^2 de 0,58), *Tempo_End* (R^2 de 0,60), *Tvr_Prog% (T)* (R^2 de 0,64). Posteriormente, os acréscimos são marginais. Portanto os dois períodos coincidem nas variáveis mais correlacionadas e mantém-se o respectivo sinal. Em geral os dias laborais apresentam menores valores dos coeficientes de determinação.

Coefficientes de Correlação *Rat_Spot%*

Quanto à variável *Rat_Spot%*, os programas com uma maior saturação publicitária e duração global e antecédidos de programas

com maior duração apresentam menores audiências comerciais. Uma maior duração desde o início do programa até à parte associada ao espaço comercial ou à inserção; maior número de partes totais de um programa ou da parte associada ao espaço comercial; maiores durações de espaços comerciais e com maior número de inserções e as maiores posições num espaço comercial são as variáveis que penalizam *Rat_Spot %*.

Uma maior duração do espaço comercial, uma maior duração global do programa, um maior número da parte associada ao espaço comercial, um maior número de inserções no espaço comercial, também está negativamente correlacionada com *Rat_Spot%*. Os ratings dos programas adjacentes e do programa favorecem as audiências comerciais, assim como os períodos com maiores níveis de *Tvr_Prog%* o que corresponde a horas mais tardias. Para este canal o valor de correlação com *Rat_Prog% (T)* é muito elevado 0,95 e com *Rat_Prog% (P)* de 0,92. Com as taxas médias para o total da TV os valores dos coeficientes de determinação também são elevados, sendo de 0,85 com o valor parcial e de 0,83 com o valor total.

A um nível intermédio constam os horários de início dos programas e o valor da fidelidade programática. Com coeficientes irrisórios constam as durações programáticas parciais, a saturação publicitária do programa anterior, a duração da inserção, a duração desde a inserção até à conclusão do espaço comercial a nível absoluto e relativo, *Ratio_Dur_Ac* e *Ratio_Início_Spot*.

➤ Modelos de Regressão Linear para *Rat_Spot%*

Para a variável *Rat_Spot%* um modelo de Regressão Linear, considerando como variável independente apenas o *rating* do programa e parcial tem um R^2 de 0,85 e com um valor de Beta de 0,925. Uma vez mais é elevada a linearidade existente entre os *ratings* programáticos e os valores comerciais. Realizando uma Regressão Stepwise seguem-se os *ratings* programáticos globais (+), *Dur_Ac* (-) e *Tvr_Prog%* (T) (-) com um R^2 de 0,89. Apenas o sinal de *Tvr_Prog%* não corresponde ao esperado, sendo negativo.

Retirando os *ratings* programáticos o modelo mais parcimonioso e com maior variância explicada engloba *Tvr_Prog%* (T) (Beta=0,91) (neste caso o valor do beta é positivo e apresenta um R^2 de 0,57), *Tempo_End_Prog* (Beta=-0,23) (R^2 de 0,66), *Tempo_Spot* (+) (R^2 de 0,70). Posteriormente, é necessário incluir inúmeras variáveis para se verificar um acréscimo de 3% no valor do R^2 .

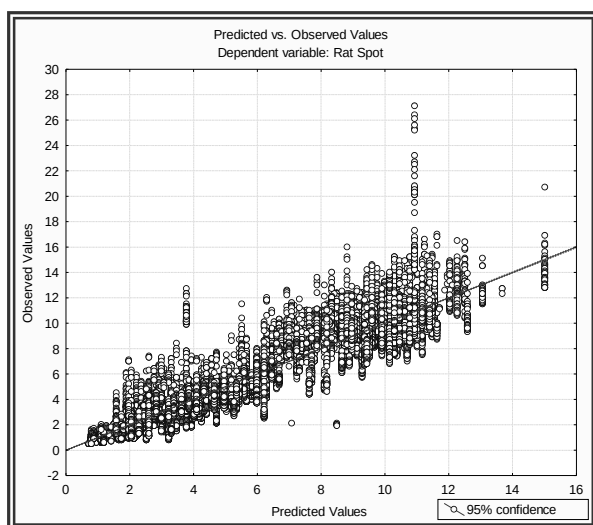


Gráfico F.7.5.2 – Valores previstos e valores observados

Na ausência do valor da taxa de audiência para o programa, a qual de antemão não será conhecida, assume uma maior relevância a taxa de audiência para o total da televisão (*Tvr_Prog%*) com uma significativa penalização ao nível da variabilidade explicada pelo modelo. Identificamos *Tvr_Prog%* como a variável mais correlacionada com os *ratings* e obviamente favorecendo-os. Também, são relevantes outras variáveis associadas aos factores temporais. Os programas com horários de conclusão mais tardios apresentam menores valores de *ratings*. Esta relação não é de todo linear como consequência das oscilações diárias verificadas nas taxas de audiências.

Se restringirmos as análises ao prime time e à noite e aos dias laborais verificamos que a variável mais relevante são os *ratings* programáticos totais (+) (R^2 de 0,30, valor muito inferior o global, mas também positivo), *Dur_Ac* (-) (R^2 de 0,40), *Rat_Prog%* (P) (+) (R^2 de 0,47), *Tvr_Prog%* (T) (-) (R^2 de 0,51 e neste caso negativo o beta), *Tempo_End_Prog* (G) (+) (R^2 de 0,58), *Tvr_Prog%* (P) (+), *Num_Part* (-), *Dur_Espaço_Comercial* (-), *Dur_Spot* (-), *Tempo_Spot* (-), *Dur_Prog* (P) (-), *Num_Promos* (-), *Dur_Prog* (T) (-) (R^2 de 0,68).

Portanto, é necessário incluir oito variáveis para se verificar um acréscimo de 8% na variabilidade explicada. Apenas o sinal de *Tvr_Prog%* (T) e de *Tempo_End_Prog* (T) não corresponde ao esperado. Os *ratings* programáticos beneficiam os comerciais e segundo estes resultados também os beneficia o horário de conclusão mais tardio. Dada a popularidade da telenovela do canal SIC referente ao início da noite este canal é o mais penalizado destacando-se as telenovelas portuguesas da TVI que acabam em horários

mais tardios. Existe, portanto, uma contra-programação não de confrontação directa, mas, sim, de transmissão de programas fortes em períodos horários mais fracos em termos de audiências por parte dos canais concorrentes.

Os valores de *Tvr_Prog%* (*T*) penalizam esta variável talvez porque nesses períodos as audiências programáticas também não atingem os seus maiores valores, sendo este canal também penalizado a nível comercial. Programas mais longos a nível global e parcial, espaços comerciais mais longos e com um maior número de promoções, inserções com maiores durações acumuladas penalizam *Rat_Spot%*. O coeficiente da variável *Tempo_Spot* também é negativo, mas a relação poderá não ser totalmente linear.

Após a conclusão das telenovelas, o que também corresponde a horas mais tardias os *ratings* comerciais obviamente são menores. Se não incluirmos os valores dos *ratings* programáticos a primeira variável a ser incluída é *Tempo_End_Prog* (+) (R^2 de 0,14), *Dur_Ac* (-) (R^2 de 0,21), *Tempo_Spot* (+), *Dur_Spot* (-), *Dur_Prog* (T) (-), *Dur_Prog* (P) (+) (R^2 de 0,38). Portanto, o coeficiente de determinação fica bastante aquém do que resulta de considerar os valores dos *ratings* programáticos. Considerando idênticos períodos horários, mas de fim-de-semana a variável mais relevante também é o *rating* programático total (+) (R^2 de 0,45, superior aos dias laborais). Seguem-se também *Dur_Ac* (-) (R^2 de 0,54, superior aos dias laborais), *Tvr_Prog%* (T) (-) (R^2 de 0,59), *Tvr_Prog%* (P) (+) (R^2 de 0,63), *Rat_Ldout%* (+) (R^2 de 0,66), *Dur_Spot* (-), *Tempo_Spot* (-), *Dur_Prog* (P) (-), *Rat_Ldin%* (+), *Num_Parties_Prog* (-), *Sat_Prog%_T* (-) (R^2 de 0,70).

Uma vez mais os *ratings* programáticos favorecem os *ratings* comerciais e o valor de *Tvr%*, mas apenas o parcial e não o total. Mantem-se o sinal dos coeficientes em relação aos dias laborais e o valor do coeficiente de determinação final também é semelhante. Se consideramos apenas as três variáveis iniciais este valor é superior ao fim-de-semana. Se excluirmos os *ratings* programáticos a variável mais relevante é *Tempo_End_Prog* (+), segue-se *Num_Promos* (-), *Dur_Fim* (-), *Dur_Spot* (-), *Tempo_Spot* (+), *Dur_Inicio_Spot* (-), *Tvr_Prog%* (T) (+) (R^2 de 0,30). Existe uma significativa penalização em relação ao coeficiente de determinação.

Ic. Resultados para o canal RTP1

Procurando avaliar a existência de interações lineares entre a variável *Spot_Ratio%* ou *Rat_Spot%* e as restantes calcularam-se os coeficientes de correlação de pearson.

➤ Coeficientes de Correlação *Spot_Ratio%*

Para a variável *Spot_Ratio%*, as inserções de espaços comerciais associados a durações programáticas parciais mais longas e inseridos em programas com uma maior duração global; inserções de maior duração e inseridos em espaços comerciais mais longos e com um maior número de inserções, ocupando as posições com uma maior duração acumulada desde o início do espaço comercial; espaço comercial de programas com maiores *ratings* face aos adjacentes e espaço comercial de programas com maior índice de saturação publicitária são os que perdem mais audiência. Portanto, todos os factores que contribuem para aumentar a durabilidade dos espaços comerciais ou da localização da inserção e do programa penaliza este índice, assim como

também a popularidade de um programa, a durabilidade e a saturação publicitária do programa anterior.

Constatou-se que, nem sempre a popularidade de um programa contribui para reter a sua audiência ao longo dos seus espaços comerciais, o suficiente por forma a aumentar este indicador. Esta variável poderá penalizar *Spot_Ratio%* uma vez que apesar dos espaços comerciais associados a programas mais populares apresentarem maiores *ratings* comerciais o incremento verificado a este nível não atenua o efeito do aumento verificado em denominador. Por isso mesmo as variáveis *Ratio_Ldout%*, *Ratio_Ldin%* e *Rat_Prog%* têm coeficientes negativos. Obviamente que o *rating* de uma inserção favorece esta variável e programas associados a horas mais tardias (no prime time e à noite) têm maiores *ratings* programáticos, mas, também, retêm mais a audiência ao longo dos espaços comerciais.

A proximidade da inserção face às posições iniciais do espaço comercial (maior *Ratio_Dur_Fim*) e do fim do programa ou do espaço comercial em que está inserido (maior *Ratio_Início_Spot*) aumenta este indicador. Este último resultado contradiz a influência negativa exercida pela variável número da parte do programa associado ao espaço comercial. Segundo as análises Anova existe uma relação directa entre os valores de *Spot_Ratio%* e o número do espaço comercial. Anteriormente tínhamos verificado que a primeira posição da inserção no espaço comercial é a que apresenta menor redução nos *ratings* comerciais (0.98 de média *Spot_Ratio*),

segue-lhe a segunda posição e a última posição. Portanto, confirma que o início e o final do espaço comercial são as posições que mais retêm a audiência. Consequentemente a duração acumulada está negativamente correlacionada com *Spot_Ratio%*.

Quanto à proximidade da inserção com o final do programa os resultados das análises Anova estão mais de acordo com o resultado associado à variável *Ratio_BS*. O quinto (0,8), o terceiro (0,895) e o sexto (0,893) intervalo e o intervalo antes do início do programa (0,87) são os que apresentam maiores valores de *Spot_Ratio%*, ou seja, os últimos espaços comerciais de programas mais longos.

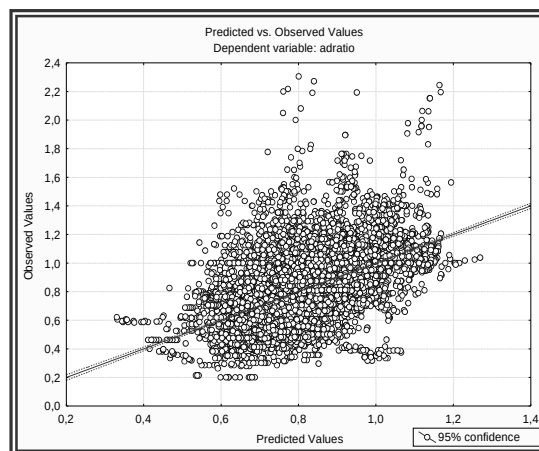


Gráfico F.7.5.3 – Valores Observados e Previstos

A maior percentagem de programas apenas têm um ou dois espaços comerciais e são os mais penalizados em termos comerciais. Portanto, a variável *Num_Parte_Prog* e *Num_Total_Partes* deveria estar positivamente correlacionada com *Spot_Ratio%*. Outros factores devem estar a influenciar esta variável.

Uma maior duração e saturação publicitária do programa anterior penaliza este índice. Em relação a *Tvr_Prog%* e ao número de promoções por espaço comercial não foi possível retirar conclusões, uma vez que os coeficientes são irrisórios. Segundo as análises *Anova*, a variável *Tvr_Prog%* é superior para alguns dos períodos com maiores valores de *Spot_Ratio%* (noite e prime time), mas para o início da tarde e para a tarde verifica-se o inverso.

A relação existente entre as variáveis não é linear. Quanto ao número de promoções por espaço comercial verificou-se que alguns dos espaços comerciais com maiores valores de *Spot_Ratio%* também maximizam o número de promoções e com menores valores de *Spot_Ratio%* também minimizam o número de promoções.

➤ Modelos de Regressão Linear para *Spot_Ratio%*

Considerando um modelo de Regressão linear para os dados comerciais do canal RTP1 com apenas as variáveis métricas e a variável *Spot_Ratio%*, o melhor modelo tem um R^2 (ajustado) de 0,25. Uma vez mais o seu valor é muito reduzido.

Os maiores coeficientes estão associados às variáveis *Tvr_Prog%* (T)/(P) (*Tvr_Prog%* (P), $Beta=-,61$ e *Tvr_Prog%* (T), $Beta=,575$). Quanto maior for a audiência total disponível maior será o valor de *Spot_Ratio%*, considerando o programa global porque se considerarmos cada uma das suas partes verifica-se o oposto (beta negativo).

Este resultado corresponde à totalidade dos dados. Se *Tvr_Prog%* (P) for superior em princípio o *rating* do programa parcial também o será, logo a variável *Spot_Ratio%* sendo a razão entre os *ratings* comerciais que costumam ser reduzidos e os valores programáticos será menor. Com *Tvr_Prog%* (T) verifica-se o oposto talvez por esta variável estar associada a um período horário mais alargado, não existindo obrigatoriamente uniformidade de *ratings* para todas as partes que o integram. Podem surgir situações em que os valores de *Spot_Ratio%* sejam superiores. Segue-se a variável *Rat_Prog%* (T) com $Beta=-,29$. Aos maiores valores de *ratings* programáticos globais corresponde menores valores de *Spot_Ratio%*, uma vez que o quociente comercial e programático será menor. Os espaços comerciais mais longos perdem mais audiência (*Dur_Espaço_Comercial*, com $Beta=-,27$), assim como as inserções precedidas por um maior horizonte temporal (*Dur_Ac*, $Beta=-,19$) e inserções inseridas em espaços comerciais associados a programas com partes mais longas (*Dur_Prog* (P), $Beta=-,11$). Os programas que se iniciem em horários mais tardios têm maiores valores de audiências programáticas e comerciais (*Hora_Início_Prog*, $Beta=,172$). Curiosamente programas com uma maior saturação publicitária (*Sat_Prog*, $beta=,19$) e programas mais longos (*Dur_Prog_Total*, $Beta=,133$) também estão positivamente correlacionados com *Spot_Ratio%*.

Obviamente os programas que apresentam maior fidelidade por parte dos telespectadores maximizam esta variável (*Fid_Prog_Total*, $Beta=,365$). O reduzido valor do R^2 associado a este modelo de Regressão denota o carácter não linear das interações existentes entre as várias variáveis consideradas, ainda que a magnitude e sinal dos declives associados a

cada uma das variáveis corresponda às tendências gerais esperadas. Restringindo as análises ao período da noite e de prime time e apenas aos dias laborais com *Dur_Prog* (P) (-) o R^2 é de 0,13. Incluindo *Dur_Ac* (-), *Tempo_Spot* (-), *Tvr_Prog%* (P) (-), *Rat_Prog%* (P) (-) e *Sat_Prog%* (-) o R^2 é de 0,48. Todas as variáveis enumeradas penalizam *Spot_Ratio%*. Ao fim-de-semana, incluindo *Sat_Prog%* (-), *Dur_Ac* (-), *Tvr_Prog%* (P) (-), *Dur_Prog* (P) (-), *Tempo_Spot* (-), *Rat_Prog%* (P) (-), *Rat_Ldout%* e *Rat_Ldin%* (+) e *Dur_Espaço_Comercial* (-) o R^2 é de 0,46. Com excepção dos *ratings* adjacentes todas as variáveis têm valores negativos dos betas.

➤ Coeficientes de Correlação para *Rat_Spot%*

Para *Rat_Spot%* a popularidade de um programa em termos absolutos e relativos e a dos programas adjacentes, os períodos de maior audiência total, programas em horários mais tardios e com uma maior duração da parte associada ao espaço comercial favorecem as audiências comerciais. Quanto maior for a fidelidade a um programa e o tempo médio despendido na duração global do programa maiores são os valores dos *ratings* comerciais. Os espaços comerciais mais longos, com um maior número de inserções e associados às partes mais avançadas de um programa; espaços comerciais de programas com uma maior durabilidade total, com um maior número de partes e com um maior índice de saturação publicitária; as posições dos espaços comerciais com uma maior duração acumulada associada quer desde o início do espaço comercial como do programa; programas antecidos por outros de maior

durabilidade e saturação publicitária são os mais penalizados ao nível dos *ratings* comerciais.

Tinhamos constatado que as posições iniciais (1º/2º), ou seja, as de menor duração acumulada são as que têm maiores valores de *ratings* e também a última, sendo as restantes mais penalizadas. Apenas a variável *Dur_Ac* apresenta um coeficiente com alguma dimensão. As variáveis *Dur_Fim*, *Ratio_Dur_Ac* e *Ratio_Dur_Fim* apresentam coeficientes irrisórios o que apenas permite depreender a respeito da ausência de inter-relações lineares. A variável *número de promoções por espaço comercial* tem um coeficiente muito reduzido, mas foi possível verificar que os espaços comerciais com cinco ou mais promoções são os que têm maiores valores de *ratings* associados. O facto de uma promoção preceder ou não uma inserção não influencia os resultados e a diferença existente entre os *ratings* associados a promoções e as restantes inserções não é muito significativa.

Comparando estes resultados com os de *Spot_Ratio%* sem dúvida que existe uma maior associação linear entre as variáveis, devendo-se em primeiro lugar seguir uma abordagem linear para o desenvolvimento dos modelos explicativos de audiências comerciais. Para *Spot_Ratio%* existe uma significativa descida nos índices de correlação associados às variáveis *ratings* de programas e valores adjacentes, hora de início do programa e quanto à taxa média para o total da televisão.

As alterações de sinal para *Spot_Ratio%* face a *Rat_Spot%* estão associadas às variáveis *Ratio_Ldin%*, *Dur_Prog (P)*, *Bsh_Prog (P)*. Os programas com maiores valores de *ratings* face ao antecessor têm menores valores de *Spot_Ratio%* e maiores valores de *ratings* comerciais. A subida nos *ratings* programáticos prevalece face à componente comercial. O mesmo sucede com a duração parcial do programa em termos absolutos e relativos. Os coeficientes em valor absoluto não excedem os 0,35 e na sua maioria são bastante reduzidos, no entanto o seu sinal corresponde quase sempre ao esperado.

➤ Modelos de Regressão Linear para *Rat_Spot%*

O melhor modelo (maior valor de R^2 e mais parcimonioso) de Regressão Linear Múltipla para a variável *Rat_Spot%* apresenta um R^2 de 0,82, incluindo as variáveis *Rat_Prog% (P)* ($\beta=0,80$), duração do espaço comercial (*Dur_Espaço comercial*, $\beta=-0,11$) e *Tvr_Prog% (P)* ($\beta=0,122$). Uma vez mais existe uma associação linear entre os *ratings* programáticos e comerciais. A duração do espaço comercial obviamente penaliza as audiências e as taxas média de audiências para o total da TV favorecem-nas. A dimensão ótima dos espaços comerciais corresponde frequentemente a seis inserções. Se incluirmos apenas o *rating* do programa associado à inserção obtemos um R^2 de 0,81 com $\beta=0,90$. Portanto a inclusão das outras duas variáveis apenas contribui para um acréscimo de 1%. Contrastando com os resultados obtidos para a variável *Spot_Ratio%*, os *rating*

comerciais parecem estar linearmente correlacionados com o *rating* do respectivo programa. Com apenas uma variável dependente obtém-se uma elevada percentagem de variabilidade explicada. Assim sendo, conseguir determinar com algum rigor os *ratings* programáticos é o ponto de partida para quantificar dentro de um certo intervalo de confiança, os *ratings* comerciais a ele associados.

Se considerarmos apenas o prime time e a noite ao fim-de-semana com *Rat_Prog% (T)* (+) o R^2 é de 0,78 e com *Rat_Ldout% (+)*, *Dur_Espaço_Comercial (-)* e *Dur_Ac (-)* alcançamos os 0,88. Portanto as audiências comerciais estão associadas a maiores valores de *ratings* programáticos, inclusive do programa seguinte e são penalizadas pela duração do espaço comercial e por uma maior duração acumulada. O coeficiente associado a *Dur_Espaço Comercial* é irrisório e o de *Dur_Ac* também, o que denota que apesar de existir uma redução nas posições intermédias talvez isto não seja de forma tão acentuada como nos canais privados onde existe um público mais jovem. Se considerarmos idênticos períodos dos dias laborais verificamos que com *Rat_Prog% (T)* o R^2 é de 0,82 (superior) e se introduzirmos também *Tempo_Spot (-)*, *Rat_Ldout% (+)*, *Rat_Ldin% (+)* e *Dur_Ac (-)* o R^2 é de 0,90. As audiências comerciais são favorecidas por maiores valores de *ratings* programáticos e dos programas adjacentes. Neste canal existe uma acentuada descida nos *ratings* programáticos no decorrer da noite, logo *Tempo_Spot* tem um coeficiente negativo e a duração acumulada também, mas o seu coeficiente é o menor.

III. Análises de Regressão Linear Múltipla Simbólicas

Para a realização de análises Simbólicas as bases de dados correspondem às que foram utilizadas nas análises de Regressão, construindo-se a partir delas objectos de segunda ordem. Nas análises Simbólicas os objectos de segunda ordem correspondem a posições de inserções (primeira, segunda, intermédias, antepenúltima, penúltima e última posições) associadas a determinado programa, dia e localização de espaço comercial (como localização define-se ser antes do início do programa ou no seu decorrer), identificados pelo seu número correspondente. Por exemplo, à inserção que corresponde ao Espaço informativo da noite_Segunda_1º Intervalo_2ª Posição. As variáveis em análise coincidem com as que foram consideradas nas análises de Regressão Linear Multivariadas, dando origem nas bases de dados simbólicas a variáveis da tipologia intervalo, diagrama, histograma, taxonómicas ou multivalor.

➤ Análises de Regressão Linear Simbólicas para a variável *Rat_Spot%*

Um modelo de Regressão Linear Simbólico para a variável *Rat_Spot%*, incluindo apenas as variáveis que podem ser conhecidas *a priori* ou calculadas ainda que com alguma margem de erro que se pretende que seja a menor possível (por exemplo, *Tvr_Prog%* que apresenta uma menor variabilidade nos seus valores) tem um coeficiente de determinação de 0,87.

Assim como se verificou para a Regressão Linear não simbólica o valor da percentagem de variância explicada é razoável. Os *ratings* comerciais são função dos valores dos *ratings* programáticos, consequentemente existe uma associação linear entre os *ratings* e os valores de *Tvr_Prog%*. Se excluirmos do modelo anterior as variáveis associadas a *Tvr_Prog%* (por exemplo, *Tvr_Prog% (P/T)*) obtemos um R^2 de 0,80 e que se traduz num decréscimo no valor de 7%.

Neste contexto as variáveis com uma maior capacidade explicativa são o período horário (variável modal, uma vez que alguns programas abrangem vários períodos dadas as oscilações existentes nos seus horários) e as horas ou tempo (hora convertida num número decimal equivalente que é da tipologia intervalo, uma vez que para ser modal a maioria dos objectos simbólicos apresentaria uma única observação por valor não sendo possível associar frequências. Para ser da tipologia multivalor também teria demasiadas observações por objecto simbólico) associado à inserção e ao início do programa, quer seja global ou apenas da parte associada ao espaço comercial.

Os coeficientes positivos estão associados a algumas tipologias programáticas de terceiro grau (o futebol, o épico, bolsa, filmes de acção, ...), aos períodos horários mais tardios (prime time, noite e pré-prime-time) e, consequentemente também à hora de início do programa (global).

Variável	Coefficiente
Prime Time	6,5
Noite	6,4
Histórico Épico	3,2
Futebol	2,4
Pre Prime time	2
Bolsas/Finanças	1,9
Casos Reais	1,84
Misto	1,4
Acção	1,15
Bsh_Prog (P)	0,19
Hora Início (G)	0,11
Hora Inicio (P)	$7,2 * 10^{-2}$
Sat_Ldin%	$2,75 * 10^{-2}$
Hora_End (P)	$2,6 * 10^{-2}$
Hora_End (G)	$2,5 * 10^{-2}$
Dur_BS	$1,05 * 10^{-3}$
Dur_Ldin	$-1,98 * 10^{-4}$
Dur_Prog (P)	$-7,9 * 10^{-4}$
Dur_Início_Parte	$-8,9 * 10^{-4}$
Dur_Espaço_Comercial	$-1,19 * 10^{-3}$
Dur_Spot	$-1,7 * 10^{-2}$
Len_Med_Spot	$-5 * 10^{-2}$
Hora_Spot	-0,12
Num_Part es_Total	-0,69
Ratio_Dur_Fim	-1,19
Ratio_Dur_Ac	-1,85
Ratio_Início_Parte	-2

Tabela F.7.5.6 – Coeficientes do modelo de RLM simbólico

Portanto, os factores temporais são determinantes. Ainda que significativos apresentam valores irrisórios as variáveis que correspondem à hora de início do programa (parcial) e de conclusão do programa (parcial e

global), ao índice de saturação publicitária do programa anterior e ao tempo que decorre desde o início do programa até ao suporte publicitário.

Com valores irrisórios e negativos consta a duração do programa anterior, a duração da parte do programa associado ao espaço comercial, a duração do espaço comercial e da inserção. Com coeficientes de maior dimensão e negativos destacam-se a hora da inserção, o número de partes que integra um programa, a duração desde a inserção até à conclusão do espaço comercial ou desde o início do espaço comercial até à inserção a nível percentual e o tempo que decorre desde o início do programa até à parte associada ao espaço comercial em termos percentuais.

Portanto, do ponto de vista linear existem maiores valores de audiências comerciais associados a determinadas tipologias (programas de grande popularidade que podem influenciar as correspondentes audiências comerciais) e a períodos horários sobretudo o prime time e a noite.

Do ponto de vista programático sem dúvida que os factores temporais são os mais relevantes e determinadas tipologias conseguem alterar os padrões de consumo da população, como, por exemplo, a transmissão de jogos de futebol ou de galas.

As restantes variáveis não se encontram correlacionadas ou apresentam coeficientes demasiado reduzidos.

Por outro lado as variáveis que penalizam em maior grau a *Rat_Spot%* correspondem à maior duração de um programa, ao qual também está associado um maior número de intervalos e partes, sendo mais desfavorecidos os espaços comerciais que mais se distanciam do início do programa (*Ratio_Início_Parte*). Uma vez que as posições intermédias apresentam menores valores e foram agregadas, destacam-se também as inserções às quais estão associadas maiores valores de *Dur_Fim* e *Dur_Ac*, uma vez que apenas as duas posições iniciais e a última são favorecidas. A hora da inserção também apresenta alguma importância, mas fica aquém dos períodos horários de prime time e noite.

Incluindo a variável *Tvr_Prog%* observou-se um acréscimo de 0,07 no coeficiente de determinação e as variáveis com uma maior capacidade explicativa correspondem a *Tvr_Prog%* (*P/T*), *Hora_Spot*, *Período* e a *Hora_Início* (*P/T*). Uma vez mais as variáveis com maiores coeficientes dizem respeito ao período horário (prime time, noite e pré-prime-time) e a determinadas tipologias como o futebol, meteorologia e os talkshows. Destaca-se também o *Tvr_Prog%* (*T*) e a duração parcial do programa. Com coeficientes negativos e de alguma dimensão, o número de partes de um programa e a duração desde o início do programa até à parte que está associada ao espaço comercial. Os programas mais longos, que contêm um maior número de partes são também os que estão associados aos períodos de menores audiências como a tarde e o final da noite ou a manhã. Se considerássemos a totalidade das variáveis obteríamos um coeficiente de determinação de 0,93.

As variáveis com uma maior capacidade explicativa são o *Rat_Prog%* (*P/T*), o *Tvr_Prog%* (*P/T*), o *Rat_Ldin%*, o período e a *tipologia de grau 2*. Portanto, uma vez mais prevalecem as variáveis associadas a períodos horários e às tipologias quando estas correspondem a programas que se diferenciam, como, por exemplo, o futebol.

Obviamente que maiores valores de *Rat_Spot%* têm que estar associados a maiores valores de *Rat_Prog%* e *Tvr_Prog%*. As variáveis com coeficientes não nulos são por ordem decrescente de valores: a *Hora_Spot*, as tipologias de transmissão de desporto (2,3), programas de apoio à comunidade (2,2), magazines (2,1), informação desportiva (1,8), informação especializada (1,5), programas religiosos (1,6), Flash sínteses (1,2), o prime time (1,04) e o pré-prime-time (1,02), a *Hora_Fim_Prog* (*G*) (0,42), o *Rat_Prog%* (*P*) (0,24), *Bsh_Prog* (*P*) (0,15), o *Rat_Ldin* (0,10). As variáveis *Num_Promos*, *Tvr_Prog%* (*T/P*), *Sat_Prog%* e *Beta_Prog* apresentam valores positivos, mas irrisórios.

Com coeficientes negativos e de alguma dimensão consta o *Ratio_Início_Parte* (-1,5) e o *Ratio_Ldin%* (-0,34). Logo, os maiores coeficientes estão associados a tipologias de programas e a períodos horários. O sinal positivo associado a *Rat_Ldin%* e o negativo associado a *Ratio_Ldin%* são curiosos. Os *ratings* programáticos ou dos programas envolventes obviamente favorecem os comerciais, logo o sinal negativo é mais difícil de entender.

Variável	Coefficiente
Magazines	0,92
Programa Religiosos	0,79
Magazines Desporto	0,41
Tempo_Início (P)	0,39
Tele-Filme	0,31
Tarde	0,10
Num_Partes_Prog	-0,10
Tempo_End (G)	-0,11
Ratio_Início_Parte	-0,26
Tempo_Início (G)	-0,42

Tabela F.7.5.7 - Variáveis com Betas de RLM simbólica significativos

➤ Modelos de Regressão Linear Simbólicos para *Spot_Ratio%*

Para a variável *Spot_Ratio%* se considerarmos apenas as variáveis conhecidas ou que podem ser calculadas obtemos um R^2 de 0,39, muito menor ao do modelo associado a *Rat_Spot%*, mas superior aos dos modelos lineares clássicos. As inserções com maiores valores de *Spot_Ratio%* correspondem a programas em horários mais tardios, mas fora dos períodos de maiores audiências caracterizados por uma audiência mais idosa e como tal com uma menor tendência para o *zapping* (*Tempo_Início (P)* e a Tarde) e a programas religiosos ou desportivos, magazines e telefilmes.

Os eventos desportivos e os programas religiosos como a Procissão das Velas ou a Missa de Bênção dos Doentes devem apresentar uma maior regularidade nos seus *ratings*, inclusive durante os espaços informativos.

Ao contrário de *Rat_Spot%* que assumia maiores valores a partir das 19H, para *Spot_Ratio%* sobressai a tarde. A tarde corresponde a um período com uma audiência que apresenta uma menor tendência para o *zapping* e durante a qual os *ratings* programáticos também são menores, consequentemente os valores de *Spot_Ratio%* são superiores. As variáveis que penalizam esta variável são o número de partes de um programa, o tempo que decorre desde o início do programa até à parte associada ao espaço comercial e o tempo de conclusão (global) e de início do programa (global).

Logo, os programas transmitidos em horários mais tardios têm menores valores (o que corresponde aos períodos de maiores audiências) e os intervalos mais afastados do início do programa também. Incluindo *Tvr_Prog%* obtemos o mesmo coeficiente de determinação.

Para um modelo de Regressão, incluindo a totalidade das variáveis obtemos um R^2 de 0,47, ligeiramente superior. As variáveis com coeficientes de alguma dimensão uma vez mais incide em tipologias e em períodos horários mais tardios. As variáveis que penalizam *Spot_Ratio%* são a duração desde o início do programa até à inserção ou à parte do programa e a hora de início global do programa.

Concluimos que as análises de Regressão Simbólicas apresentam maiores valores de coeficientes de determinação face aos modelos multivariados, especialmente os relacionados com a variável *Spot_Ratio%*. Uma vez mais para os índices comerciais existem relações de linearidade face aos valores programáticos. As variáveis temporais ou determinadas tipologias são as que assumem uma maior relevância.

IV. Análises de Regressão Não Linear

➤ Resultados para o Canal RTP1 - Modelos de Regressão Não Linear para *Rat_Spot%*

Por aplicação do algoritmo CART resulta uma percentagem de 94% de variância explicada para a amostra de teste que representa 30% das observações. Os *ratings* comerciais estão associados aos *ratings* programáticos globais. A primeira regra traduz-se em *se Rat_Prog% (T) > 6,75%* então *Rat_Spot% Medio = 8,3%* e *Rat_Spot% Medio = 2,7%* caso contrário. Corresponde a um diferencial de 5,6% entre os dois conjuntos de observações.

Para os valores mais elevados dos *ratings* programáticos globais os valores de *Rat_Spot%* dependem das audiências programáticas parciais (*se Rat_Prog (P) > 13%*, *Rat_Spot% Medio = 14%* e *7,6%* caso contrário (*Dif. de 6,4%*)). Para os maiores valores de *ratings* programáticos parciais quanto maior for o *rating* da parte sucessora maiores são as audiências comerciais (*se Rat_Ldout (P) > 17,7%*, *Rat_Spot%_Medio = 17,6%* e *12,3%* caso contrário (*Dif 5,3%*)).

Atendendo a estes resultados as audiências comerciais dependem dos valores dos *ratings* programáticos globais, parciais e adjacentes. A diferença entre os valores médios dos *ratings* comerciais segundo os valores programáticos globais é muito acentuada (5,6%). Do ponto de vista programático verificamos que existem desníveis muito significativos, condicionando o desempenho comercial dos programas.

Os períodos de maiores audiências e os programas desportivos têm uma repercussão muito significativa nos valores comerciais. Em relação aos valores adjacentes ao espaço comercial em análise em primeiro lugar dependem da parte que os antecede (*Rat_Prog% (P)*) e posteriormente da popularidade do programa ou da parte que se inicia após a sua conclusão. Por exemplo, no pré-prime-time os espaços comerciais podem ser favorecidos por maiores valores de *ratings* de lead out que estão associados a uma maior proximidade face ao espaço informativo da noite.

Para valores de *ratings* parciais inferiores aos 13% (em princípio neste caso excluem-se a maior parte dos programas desportivos ou outros de maior popularidade e também o espaço informativo da noite). os *ratings* comerciais são superiores se estiverem associados aos períodos de pré-prime-time e noite (*Rat_Spot Medio = 8,5%*, *5,3%* c.c., *Dif 3,2%*). O diferencial de valores não é tão significativo como quando se verificam maiores valores de *ratings* programáticos.

Estes períodos correspondem aos períodos adjacentes aos que apresentam maiores audiências do ponto de vista programático, quando não são introduzidos programas de maior popularidade.

Para os menores valores dos *ratings* programáticos globais (*Rat_Prog%* (T) <6,75%) os *ratings* comerciais também são função das audiências programáticas parciais (se *Rat_Prog%* (P) >4%, *Rat_Spot% Medio*=4% e 2% c.c., Dif 2%). No entanto, o diferencial de valores é de apenas 2%. Para os maiores valores de audiências programáticas parciais quanto menor for a duração desde o início do programa até à inserção publicitária mais estes suportes são beneficiados. O *split* é realizado nos 25 segundos. Para valores inferiores a este valor *Rat_Spot% Medio*=5,6% ou 4% c.c., Dif 1,6%). Para valores de *ratings* parciais inferiores aos 4%, as audiências comerciais são uma vez mais beneficiadas por maiores valores de *ratings* globais.

Portanto, para os menores valores de *ratings* programáticos a proximidade da inserção face ao início do programa poderá favorecer o suporte, o que corresponderá a períodos com tendências descendentes nos valores dos *ratings*. Para os valores mais reduzidos onde podem existir patamares de valores, as audiências comerciais são favorecidas por maiores audiências programáticas globais.

Se excluirmos das análises anteriormente realizadas os *ratings* programáticos globais e os *ratings* parciais do programa associado ao spot e adjacentes verificamos que em primeiro

lugar os níveis de audiências comerciais são função de *Tvr_Prog%* (T). Apresenta 85% de variância explicada o que corresponde a uma penalização de 9% face ao modelo anterior. Uma vez mais são os valores globais e não os parciais que mais condicionam as audiências comerciais. A variável que traduz as audiências médias para o total da TV será a melhor substituta dos valores dos *ratings* programáticos, dada a correlação linear existente entre as duas variáveis. Para valores de *Tvr_Prog%* (T) >31%, *Rat_Spot% Medio*=8,6% e 3,2% c.c. (Dif. 5,4%).

Neste caso os *shares* parciais condicionam os *ratings* (se *Shr_Prog%* (P) >38%, *Rat_Spot% Medio*=14% e 7,8% c.c., Dif 6,2%). Poderá corresponder aos eventos desportivos que apresentam elevados valores de taxas médias para o total da TV, sendo os valores comerciais tanto superiores quanto maiores forem os valores de *shares* destes eventos. Para valores de *shares* inferiores aos 38% as taxas de audiências comerciais médias são muito menores e certamente correspondem a programas dos períodos de maiores audiências.

Para valores de *Tvr_Prog%* (T) <31% o que condiciona os *ratings* são os períodos horários com os maiores valores associados ao almoço, Pré-Prime-time, Prime Time e Noite (*Rat_Spot Medio*=4,6% e 2% c.c.). O diferencial de valores é muito menor, apenas de 1,4%. Correspondem aos mesmos períodos que maximizam as audiências programáticas. Para este subconjunto de faixas horárias obviamente que se destaca o prime time (*Rat_Spot Medio* =10% e 4% c.c., Dif. 6%), que corresponde ao espaço informativo da noite o qual apresenta grandes valores de fidelidade por parte do

público mais assíduo deste canal. Portanto, os *ratings* comerciais são essencialmente função dos *ratings* programáticos e dos factores temporais. Considerando a globalidade dos dados, os factores de planeamento comercial não são as variáveis mais relevantes.

Excluindo os valores de *Tvr_Prog%* e de *Shr_Prog%* o modelo apresenta 80% de variância explicada e que corresponde a uma penalização em 5% face ao modelo anterior. As audiências são função do período horário. Para os períodos de pré-prime-time e prime time o que condiciona os *ratings* é a hora da inserção sendo superiores para *Tempo_Spot* > 0,81.

Neste caso os suportes diferenciam-se em função dos programas com três eventos desportivos a apresentar *Rat_Spot% Medio*=18% e 9% para os restantes programas (destes por sua vez os *ratings* maximizam-se para o telejornal (*Rat_Spot Medio*=10% e 7,6% c.c.). Portanto, as audiências são função de programas específicos, destacando-se os eventos desportivos e o espaço informativo da noite.

Para *Tempo_Spot* < 0,81, mas dentro do pré-prime-time a duração global do programa condiciona os valores dos *ratings* (se *Dur_Prog* > 7270s, *Rat_Spot% Medio*=11,3%; 4,7% c.c.). Portanto, corresponderá a programas de fim-de-semana que também maximizam os valores dos *ratings* programáticos.

Os seguintes períodos com maiores valores de *ratings* são o almoço e a noite (*Rat_Spot% Medio*=4,5%), sobressaindo os noticiários (almoço), o futebol (noite) e os filmes de romance e os musicais (noite).

Para os períodos da tarde e da madrugada que também minimizam os *ratings* programáticos os valores comerciais (*Rat_Spot_Med*=2%) estão condicionados pela hora de início do programa parcial. Portanto, os períodos horários e determinados programas condicionam as audiências programáticas. A hora influencia os resultados no pré-prime-time, tarde e madrugada.

Se excluirmos a variável referente aos períodos horários no primeiro nível da Árvore surge a variável de identificação do programa com os dois espaços informativos, os programas desportivos e outros programas de prime time e do início da noite a maximizarem os índices de audiências. Para estes programas os que apresentam uma maior duração global (portanto, os eventos desportivos) maximizam as audiências e neste caso as partes programáticas com início após as 19.30H.

Existe uma selecção programática fortemente condicionada por restrições temporais. Para os programas de menor duração global sobressaem os espaços informativos da noite. Para o segundo subconjunto de programas definido no primeiro nível da Árvore será a hora da inserção comercial que os condiciona, maximizando-se os valores de *ratings* a partir das 18.30H.

Portanto, as variáveis que condicionam os *ratings* são essencialmente os factores temporais e o facto de estarem inseridos em programas específicos. Os factores de planeamento para este canal não são relevantes considerando a totalidade da emissão diária, por se verificar uma selecção programática mais pontual. Os modelos de previsão de audiências que recorrem a séries temporais parecem ser uma boa opção para este canal, sendo necessário ajustar os resultados quando se transmitem determinados eventos.

Met.	Dados	% de Variância Explicada	Variáveis
CAR T	RTP1	94%	Rat_Prog% (T) Rat_Prog % (P) Rat_Prog_Ldout % (P) Período Dur_BS
CAR T	RTP1	85%	Tvr_Prog% (T) Shr_Prog% (P) Período
<i>CART</i>	RTP1	80%	Período Tempo_Spot Id_Programa Hora_Prog

Tabela F.7.5.8 – Resultados das análises de Regressão Linear para o canal RTP1

➤ Resultados para o Canal RTP1 - Modelos de Regressão Não Linear para *Spot_Ratio%*

Assim como sucedeu nos modelos lineares a percentagem de variância explicada fica muito aquém à calculada para os modelos que visam a caracterização da variável *Rat_Spot%*. A variável utilizada no primeiro nível da Árvore corresponde ao quociente entre o valor do *rating* programático parcial e o valor de *rating* de *lead out*. Se este valor for superior à unidade, ou seja, $Rat\_Prog\% > Rat\_Ldout\%$ o valor de $Spot\_Ratio\%_{Médio} = 0,73$ e $Spot\_Ratio\%_{Médio} = 0,88$ caso contrário. Portanto, se os *ratings* programáticos de programas sucessores forem superiores aos do programa ($Rat\%/Rat\_ldout\% < 1$) também o serão os índices comerciais em termos relativos. Em períodos com tendências ascendentes nos *ratings* a crescente presença de novos elementos do painel poderá contribuir para uma menor redução dos valores de *Spot_Ratio%*.

Quando os valores programáticos são superiores aos valores de programas sucessores verifica-se o inverso. Para estes programas ($Rat\%/Rat\_ldout\% < 1$) os espaços comerciais beneficiam quando estão associados aos períodos da tarde, prime time e noite (média de *Spot_Ratio%* de 0,96 e 0,70 c.c.). Quanto mais popular for o programa seguinte mais o espaço comercial será beneficiando favorecendo o indicador *Spot_Ratio%*. Curiosamente não inclui o período de pré-prime-time que apresenta uma tendência crescente de *ratings*, podendo os suportes comerciais ser pontualmente favorecidos por uma maior proximidade face ao espaço informativo.

Inclui a tarde que também apresenta uma tendência ascendente de valores de *ratings* (14.30H-19H), o prime time associado ao espaço informativo em que a segunda parte em geral maximiza as audiências e o período da noite.

Segundo esta Árvore os períodos de pré-prime-time e da madrugada minimizam os valores de *Spot_Ratio%*. No pré-prime-time talvez a redução seja devido ao público aproveitar os espaços comerciais para realizar outras actividades, apesar da tendência ascendente de *ratings* programáticos que se verifica neste período. Para o conjunto de períodos que maximizam os valores de *Spot_Ratio%* (Tarde, Prime Time e Noite) estes serão tanto maiores quanto menor for o quociente entre *Rat_Prog%* e *Rat_Ldin%* (se $\frac{Rat\_Prog\%}{Rat\_Ldin\%} < 0,50$, *Spot_Ratio%_Médio*=1,18 ou 0,93 caso contrário).

Portanto, o indicador *Spot_Ratio%* é beneficiado por valores de *ratings* programáticos inferiores a metade do valor que os antecede, o que corresponderá ao período da noite. Para os maiores valores de $\frac{Rat\_Prog\%}{Rat\_Ldin\%} (>1)$ e na tarde, prime time e noite o valor de *Spot_Ratio%* será tanto maior quanto maior for o valor de fidelidade global do programa (se *Fid_Prog%* (T)>18%, *Spot_Ratio%_Médio*=0,80 e 0,63 caso contrário) o que deverá corresponder ao espaço informativo. Portanto, quanto maior forem as audiências da parte programática que lhe está associada face ao valor sucessor os valores de *Spot_Ratio%* beneficiam com a fidelidade que é dirigida ao programa.

Portanto, os valores das audiências relativas são função da popularidade do programa sucessor, o qual exercendo uma maior atractividade sobre o público atenua a tendência ao *zapping* mantendo uma maior estabilidade das audiências no decorrer dos espaços comerciais que o antecedem. Os factores temporais são uma vez mais relevantes, maximizando-se este indicador para a tarde, prime time e noite.

Se excluirmos os valores de *ratings* de *lead out* em termos absolutos e relativos, a variável utilizada no primeiro *split* é a duração acumulada desde o início do espaço comercial até à transmissão da inserção comercial ($Dur\_Ac < 20$ Seg, *Spot_Ratio%_Médio*=1 e 0,78 caso contrário). Portanto, as posições mais próximas do início do espaço comercial maximizam estes valores mantendo-se aos níveis das audiências programáticas que lhe estão associadas. No entanto, este valor é muito reduzido, apenas 20 segundos, após o qual se verifica um considerável decréscimo. Para estas inserções destacam-se as partes programáticas com valores superiores aos antecedentes ($\frac{Rat\%}{Rat\_Ldin\%} > 1$) e que marcam o final do programa, estando o espaço comercial associado localizado entre programas (se *Loc_Espaço_Comercial*=Entre, *Spot_Ratio%_Médio*=1,17 e 0,90 c.c.). Para as partes localizadas no decorrer de um programa se os seus valores forem inferiores aos antecedentes a sua capacidade para maximizar as audiências para as inserções iniciais é menor (0,90 de Média para *Spot_Ratio%*). Para as durações acumuladas superiores aos 20 Segundos (*Spot_Ratio%_Médio*=0,78) os períodos horários condicionam estes valores.

O pré-prime-time e a madrugada são os mais penalizados (*Spot_Ratio%_Médio*=0,66 e 0,80 caso contrário). Neste contexto beneficiam as inserções associados às menores durações programáticas parciais (se *Dur_Prog* (P)<1151s, *Spot_Ratio%_Médio*=0,88 e 0,66 caso contrário) e neste caso as séries, concursos e os resumos de desporto (os filmes e os *reality shows* são as tipologias mais penalizadas).

Para as maiores durações parciais destacam-se os espaços informativos e os eventos desportivos. Para os restantes períodos (Tarde, Principio Tarde, Noite e Prime Time) e se *Dur_Ac*>20 seg, se o *rating* programático parcial superar o global (*Spot_Ratio Médio*=0,77 e 0,86 c.c.) os espaços comerciais mais beneficiados são os que se localizam entre programas ou no terceiro intervalo (*Spot_Ratio Médio*=0,96, e 0,65 c.c.). Os espaços comerciais associados às partes finais de um programa são os mais beneficiados em termos relativos e os primeiros e segundos espaços comerciais são mais penalizados. Se os *ratings* parciais forem inferiores aos globais (*Spot_Ratio%_Médio*=0,86) os valores relativos das audiências comerciais são superiores para *Tempo_Início_Prog* (P)<0,89 (*Spot_Ratio%_Médio*=0,88 e 0,73 caso contrário). Atendendo às audiências comerciais em termos absolutos as variáveis que as condicionam para o canal estatal são muito semelhantes às identificadas para as audiências programáticas, incidindo essencialmente em períodos horários ou em variáveis relacionadas, programas em concreto e em tipologias de programas.

Para os valores de *Spot_Ratio%* que traduzem os valores relativos das audiências comerciais assumem uma maior importância para além dos factores temporais também algumas variáveis de planeamento, como, por exemplo, as localizações dos espaços comerciais e as durações acumuladas até ao suporte publicitário.

Met.	Dados	% Exp.	Variáveis
CART	RTP1	53%	Rat_Prog/Rat_Ldout
			Período
			Rat_Prog/Rat_Ldin
CART	RTP1	52%	Fidelidade
			Dur_Ac
			Localização_espaço comercial
			Rat_Prog/Rat_Ldin
			Período
			Tipologia
			Tempo_Início_Prog

Tabela F.7.5.9 – Resultados das análises de Regressão Linear para canal RTP1 e para a variável *Spot_Ratio%*

➤ Modelos de Regressão Não Linear para *Rat_Spot%* – Resultados para o canal SIC

Prosseguindo as análises com os dados do canal SIC e para a variável *Rat_Spot%* a variância explicada é de 90% e verificamos que a variável utilizada no primeiro nível da Árvore é *Rat_Prog%* (P) (se *Rat_Prog %*

(P)>8,7%, *Rat_Spot%_Medio*=9,2 e 3,9% caso contrário). Para os maiores valores de *ratings* programáticos parciais, as audiências comerciais subdividem-se uma vez mais segundo os valores desta variável. Se estes valores forem superiores aos 13% e se o programa for antecedido também por maiores valores de *ratings* maximizam-se as audiências comerciais e para valores inferiores aos 13% será o período horário que influencia os resultados, sendo estes superiores para os espaços comerciais inseridos em programas com início às 20H, 21H e 22H durante os dias laborais e ao fim-de-semana para os mesmo horários agregando-se também o horário das 13H (*Rat_Spot_Médio*=9% e 6% c.c.).

Os períodos que minimizam os *ratings* são as 19H, as 23H e as 24H durante os dias laborais e as 14H, 15H, 19H, 23H ao fim-de-semana. Portanto, para os maiores valores de audiências programáticas parciais e superiores aos 8,7% e inferiores aos 13% as audiências comerciais maximizam-se nos períodos onde existe uma maior estabilidade nos valores dos *ratings*, ou seja, entre as 20H e os programas com início entre as 22H e as 23H.

Para os menores valores de audiências programáticas parciais e inferiores aos 8%, *Rat_Spot%* é função das audiências globais. Verificamos que uma maior associação entre os *ratings* parciais e os *ratings* globais corresponde a intervalos extremos de valores de *Tvr_Prog%*.

Quando *Rat_Prog%* (*T*) for superior aos 4% maximizam-se entre as 19H e as 22H em dias laborais e também entre as 13H e as 14H ao fim-de-semana. Portanto, para o canal SIC os *ratings* comerciais são função das audiências programáticas parciais e globais e de factores temporais, sendo estes maximizados nos períodos de prime time e no horário associado à principal telenovela da noite. A popularidade de um programa incentiva à permanência das audiências no decorrer dos seus espaços comerciais.

Se excluirmos estas variáveis são os níveis de *Tvr_Prog%* (*P*) que condicionam os *ratings* comerciais. Para valores de *Tvr_Prog%* inferiores aos 30%, os *ratings* comerciais maximizam-se no prime time, noite e à hora do almoço e minimizam-se durante a tarde e madrugada. Para os períodos de maiores audiências os *ratings* são função de determinados programas e para a tarde e madrugada são uma vez mais função dos níveis de *Tvr_Prog%*.

Portanto, para os menores valores de *Tvr_Prog%*, os factores temporais condicionam os *ratings* comerciais. Para valores de *Tvr_Prog%* (*P*)>30% e para os programas com maiores valores de fidelidade (*Fid_Prog%*>43%) a variável *Rat_Spot%* é superior quando a duração desde o início do programa até à parte associada ao espaço comercial for menor. Para valores de fidelidade inferiores aos 43% destacam-se com maiores valores de *Rat_Spot%* determinados programas.

Se excluirmos os nomes de programas, assim como os valores de fidelidade e de *Tvr_Prog%* surgem como sendo mais relevantes as variáveis temporais (horas, tempos, períodos ...). No primeiro nível consta *Tempo_Spot* (se $Tempo\_Spot > 0,81$, $Rat\_Spot\%\_Médio=8,5\%$, 3,6% caso contrário). Os *ratings* maximizam-se com $Tempo\_Spot > 0,81$ e $Tempo\_Spot < 0,97$ ($Rat\_Spot\%\_Médio=9,1\%$) e neste caso no período semanal ($Rat\_Spot\%\_Médio=9,8\%$, 7,6% caso contrário).

Para os períodos de maiores audiências programáticas os *ratings* comerciais são superiores em dias laborais, em detrimento do fim-de-semana. Durante os dias laborais as tipologias de telenovela, humor e as galas apresentam os maiores valores de *ratings* e são inferiores para programas de notícias, programas de apoio à comunidade e concursos. Para $Tempo\_Spot < 0,81$ as variáveis temporais condicionam os resultados, maximizando-se no pré-prime-time e no início da tarde. Neste caso para *Hora_Spot* anterior às 13.30H, o que corresponde ao espaço informativo da tarde.

Portanto, as audiências comerciais são uma vez mais função de factores temporais e de programas em concerto de maior popularidade. Também, para o canal privado SIC os modelos de séries temporais parecem ser uma opção a considerar, uma vez que as audiências estão fortemente condicionadas por factores temporais. A influência dos eventos desportivos em contra-programação ou transmitida no próprio canal deve ser considerada.

Met.	Dados	% Exp.	Variáveis
CART	SIC	90%	Rat_Prog (T) Período Rat_Prog(P)

Tabela F.7.5.10 – Resultados das análises de Regressão não linear para o canal SIC

➤ Análises de Regressão Não Linear para *Spot_Ratio%* – Resultados para o canal SIC

Conduzindo idênticas análises com a variável *Spot_Ratio%* uma vez mais a percentagem de variância explicada não alcança os 50%. Os valores de *Spot_Ratio%* são função em primeiro lugar das tipologias de Grau 2. O indicador maximiza-se para as telenovelas, noticiários, magazines e passatempos e minimizam-se ($Spot\_Ratio\%\_Médio=0,88$ e 0,73 c.c.) para os filmes, séries, concursos, talk-shows, galas, documentários e cerimónias de desporto. Para as tipologias mais penalizadas os espaços comerciais entre programas maximizam os valores de *Spot_Ratio%* ($Spot\_Ratio\%\_Médio=0,85$ e 0,7 caso contrário) e neste caso os espaços comerciais mais longos e com mais de 30 inserções. Para os espaços comerciais mais breves os valores de *Spot_Ratio%* são superiores quando associados ao final de um programa com *ratings* globais inferiores aos 8% ($Spot\_Ratio\%\_Médio=1$ e 0,7 c.c.). Para os espaços comerciais no decorrer dos programas os valores de *Spot_Ratio%* maximizam-se para a maior parte dos espaços destinados à transmissão de filmes ($Spot\_Ratio\%\_Médio=0,80$ e 0,60 caso contrário).

Para as tipologias que apresentam maiores valores de *ratings* os espaços comerciais associados às partes que terminam antes de *Tempo_End* ($P=0,79$ (*Spot_Ratio%_Médio*=0,95 e 0,83 caso contrário) são os mais beneficiados. Assim como se verificou para o canal estatal RTP1 e para *Spot_Ratio%* as variáveis de planeamento comercial assumem alguma relevância e em especial a localização do espaço comercial e o número de inserções. Os factores temporais e a popularidade dos programas também influenciam este indicador.

➤ **Análises de Regressão Não Linear para *Rat_Spot%* – Resultados para o canal SIC (Prime time e noite)**

Se restringirmos as análises ao prime time e à noite verificamos que *Rat_Spot%* é função das audiências programáticas globais (se *Rat_Prog%* (T) $>11\%$, *Rat_Spot%_Médio*=10,7%, 7% caso contrário). Para valores superiores aos 11%, as audiências comerciais são superiores quando as inserções estão associadas à principal telenovela da noite, a programas de humor da noite e as programas de muito curta duração intercalados entre eles (*Rat_Spot%_Médio*=12%). As audiências comerciais são essencialmente superiores para os programas sucessores do espaço informativo da noite e para a principal telenovela da noite na maior parte do período em análise.

Em alguns anos a SIC, por vezes, adoptou por transmitir apenas uma telenovela brasileira e apostar na ficção nacional, coincidindo o horário de conclusão da principal telenovela com a hora a partir da qual se verifica um

notório decréscimo nos valores das audiências programáticas. Intercalam entre o espaço informativo da noite e a telenovela da noite, alguns programas que resultem apelativos à maioria da audiência, mas que simultaneamente são menos onerosos, como, por exemplo, as produções nacionais de humor.

Para valores de *Rat_Prog%* (T) $<11\%$ uma vez mais os resultados diferenciam-se em função de programas específicos. Para estes programas verificamos que os espaços comerciais antes do seu início apresentam maiores valores do que os que estão localizados no seu decorrer (*Rat_Spot%_Médio*=9%, 7% caso contrário). Para os espaços comerciais no seu decorrer as audiências são tanto superiores quanto maiores forem as audiências programáticas globais. Para os restantes programas as audiências comerciais também são função dos *ratings* globais. Para os programas transmitidos entre as 20H e as 24H verificamos que as audiências programáticas globais condicionam fortemente as audiências comerciais. Para os maiores valores de *Rat_Prog%* verificamos que *Rat_Spot%* maximiza-se para determinados programas que coincidem com os programas sucessores ao espaço informativo da noite, sendo ligeiramente inferiores para os espaços informativos e alguns programas da noite assim como de fim-de-semana.

Para os programas com menor valor de *Rat_Prog%* uma vez mais os resultados diferenciam-se para subconjuntos de programas, influenciando uma vez mais os resultados os respectivos *ratings* programáticos e a localização do espaço comercial.

Para o segundo subconjunto de programas que apresenta maiores valores de *ratings* programáticos verificamos que os espaços comerciais antes do seu início e não no seu decorrer apresentam maiores valores de *ratings* comerciais. Portanto, são as audiências globais e não as parciais que condicionam os *ratings* comerciais. Apenas quando um programa consegue sustentar os níveis de audiências no seu decorrer podem os seus espaços comerciais ser beneficiados.

Se excluirmos das análises a identificação dos programas, os *ratings* estão condicionados pelas audiências programáticas globais. Para os maiores valores de *Rat_Prog%* beneficiam os programas com maiores valores de *Rat_Ldin%*. Para valores de *Rat_Ldin%* superiores aos 10% *Rat_Spot%* será tanto maior quanto maior forem os seus níveis de *Tvr_Prog%*. Portanto, os programas mais populares e anteceditos por outros também com maiores valores de *ratings* saem mais beneficiados quanto maior forem as taxas média de audiência para o total da TV. Para os programas que apresentam um valor de *Rat_Prog%* inferior aos 11%, a audiência parcial é a variável que condiciona o desempenho dos seus espaços comerciais. Para *Rat_Prog%* (P)>8%, *Rat_Spot%_Médio*=8% e 5,6% caso contrário. Para os maiores valores de audiências programáticas parciais são os mais favorecidos os espaços comerciais com um número de inserções inferior a 21 (*Rat_Spot%_Médio*=10% e 7,6% caso contrário).

Para os espaços comerciais com um maior número de inserções será a hora da inserção que influencia os resultados, sendo estes superiores para uma Hora_Spot anterior às 22.30H (*Rat_Spot%_Médio* =8% e 6,6% c.c.).

Para *Rat_Prog%* (P) <8% será também a hora da inserção que os condiciona (se *Tempo_Spot* < 0,98, ou seja, anterior a 23.30H, *Rat_Spot%_Médio*=6,3% e 4,6% c.c.). Portanto, as audiências comerciais são em primeiro lugar função das audiências programáticas globais que quando superiores condicionam *Rat_Spot%* através dos níveis de *Tvr_Prog%* e da popularidade do programa antecessor. Quando as audiências globais são menores verificamos que os valores parciais são os que os condicionam. Quanto maior for a popularidade das partes programáticas mais estas beneficiam os seus espaços comerciais no caso destes apresentarem uma duração mais reduzida. Caso contrário a hora em que os espaços comerciais são transmitidos é determinante para o seu desempenho. Quando as audiências parciais são menores os factores temporais influenciam os resultados.

➤ Modelos de Regressão Não Linear para *Rat_Spot%* – Resultados para o Canal TVI

Por último para o canal TVI as audiências comerciais são função em primeiro lugar das audiências programáticas parciais e não globais, como sucedeu com a SIC. Para valores inferiores aos 4,45% as audiências médias para o total da TV condicionam os resultados (se *Rat_Prog%* (T)>2,5%, *Rat_Spot%_Médio*=3,3% e 1,8% caso contrário). Para valores parciais de *ratings* entre os 4,45% e os 7,7%, obtemos 4,65% de *Rat_Spot%_Médio*.

Se $Rat\_Prog\% (P) < 11,65\%$, a variável que identifica os programas condiciona os *ratings* sendo estes máximos para programas de prime time e da noite. Para valores de *ratings* parciais superiores aos 11,65% as audiências globais condicionam as inserções. Portanto, as audiências comerciais estão fortemente relacionadas com as audiências programáticas parciais e globais. Para os intervalos mais extremos de valores programáticos parciais quanto maior for a capacidade de um programa para sustentar as audiências no seu decorrer mais favorecidos também são os respectivos espaços comerciais. Os espaços comerciais associados a programas que estão inseridos em períodos de maiores audiências, mas, também, de grande popularidade obviamente que são mais beneficiados.

Excluindo das análises os valores de *ratings* e a identificação dos programas o que condiciona os *ratings* comerciais é a hora da inserção (se $Hora\_Spot > 19.30H$ $Rat\_Spot\% \text{ Médio} = 9,56\%$, e 3,5% caso contrário). A partir do pré-prime-time quanto menor for a duração acumulada desde o início do espaço comercial até à transmissão da inserção maiores são os *ratings* (se $Dur\_Ac > 56$ seg., $Rat\_Spot\% \text{ Médio} = 11,2\%$ e 9,25% caso contrário).

Para uma $Dur\_Ac > 56$ seg. influência os resultados os valores de fidelidade globais. Para valores de fidelidade superiores aos 31%, $Rat\_Spot\% \text{ Médio} = 9,6\%$ e 7% caso contrário. Para as inserções transmitidas antes das 19.30H, os níveis de $Tvr\_Prog\% (T)$ influenciam os resultados (se $Tvr\_Prog\% (T) > 15,5\%$, $Rat\_Spot\% \text{ M.edio} = 5,1\%$ e 2,64% caso contrário).

Para níveis de $Tvr\_Prog\% > 15,5\%$ as audiências maximizam-se para a Taça de Portugal ($Rat\_Spot\% \text{ Médio} = 51,5\%$ e 5% para os restantes programas). Portanto, uma vez mais os factores temporais são determinantes. Nos períodos de maiores audiências as posições iniciais são as mais beneficiadas e as restantes em função dos valores de fidelidade ao programa. Para os restantes períodos horários os valores de $Tvr\_Prog\%$ condicionam os resultados ou a transmissão de algum programa de grande popularidade, como, por exemplo, um evento desportivo que maximizará não só as audiências programáticas, como também as comerciais.

➤ Modelos de Regressão Não Linear para $Spot\_Ratio\%$ – Resultados para o canal TVI

Para a variável $Spot\_Ratio\%$, a Árvore apresenta uma variância explicada no valor de 48% e a variável que mais condiciona este indicador é o valor de $Tvr\_Prog\% (T)$. Para valores superiores aos 5% ($Spot\_Ratio\% \text{ Médio} = 0,85$ e 1,17 caso contrário) os valores de $Spot\_Ratio\% \text{ Médios}$ são mais reduzidos uma vez que os *ratings* programáticos em denominador incrementam o seu valor. Para os maiores valores de $Tvr\_Prog\% (T)$ a duração acumulada condiciona $Spot\_Ratio\%$ (se $Dur\_Ac < 63$ seg., $Spot\_Ratio\% \text{ Médio} = 0,99$ e 0,83 caso contrário).

As posições iniciais, até um máximo de 63 segundos, são as que mais se aproximam dos valores dos *ratings* programáticos. Para as restantes posições as audiências do programa anterior também penalizam esta variável.

Se *Rat_Ldin%* superar os 3%, *Spot_Ratio%* maximiza-se para as durações programáticas globais inferiores aos 2614 seg. Para valores de *Rat_Ldin%* inferiores a 3% os níveis de *Tvr_Prog%* penalizam esta variável. Para valores de *Tvr_Prog%* inferiores aos 5% (1,17 de Média de *Spot_Ratio%*), o que condiciona esta variável são os valores de *Dur_Inicio_Spot*.

Para valores inferiores aos 700 seg. *Spot_Ratio%* é superior (*Spot_Ratio%_Médio*=1,35 e 0,77 caso contrário) e neste caso o valor de *Dur_Prog* (T) se superior aos 1777 seg. favorece *Spot_Ratio%* (*Spot_Ratio%_Médio*=1,57 e 1 caso contrário). Neste caso se *Ratio_Dur_Fim* >0,78, *Spot_Ratio%_Médio*=2 e 1,46 caso contrário. A variável *Ratio_Dur_Fim* representa a percentagem da duração para o final do espaço comercial a partir da inserção.

Nesta Árvore surge um maior número de variáveis associadas a factores de planeamento comerciais. Os níveis de *Tvr_Prog%* penalizam esta variável, uma vez que o acréscimo nos *ratings* programáticos não é acompanhado proporcionalmente a nível comercial. As durações acumuladas e programáticas penalizam *Spot_Ratio%*, assim como os *ratings* programáticos, sejam eles dos programas ou de *lead_In*. As maiores durações globais podem incentivar a audiência a utilizar os espaços comerciais para realizar outras actividades. Os espaços comerciais mais beneficiados são os que estão inseridos em períodos com menores audiências médias para o total da TV e consequentemente programáticas (*Spot_Ratio_Médio*=1,17).

Neste caso a maior proximidade da inserção face ao início dos programas favorece *Spot_Ratio%* (*Spot_Ratio%_Médio*=1,35) em especial para os programas de maior duração global (>1777 seg.) (*Spot_Ratio%_Médio*=1,57) e desde que correspondam no máximo até 22% da duração do espaço comercial (*Spot_Ratio%_Médio*=2). Portanto nos períodos de menores audiências por vezes os valores de *Rat_Spot%* podem superar os valores de *Rat_Prog%* (P) o que também se deve às durações envolvidas.

➤ Modelos de Regressão Não Linear para *Spot_Ratio%* (a partir prime time) – Resultados para o canal TVI

Se restringirmos as análises aos períodos de prime time e à noite os *ratings* comerciais estão condicionados pelas audiências programáticas para o total da TV. Quando estes valores forem superiores aos 12%, as durações parciais condicionam os resultados com valores superiores aos 1000 seg. a favorecem *Rat_Spot%*. Portanto, por si só as maiores durações programáticas não penalizam as audiências comerciais. Para as menores durações o período da noite supera o prime time e para as restantes durações os valores de *ratings* programáticos e de *Tvr_Prog%* favorecem a variável *Rat_Spot%*.

Apenas para valores de *Tvr_Prog%* (P)<23% os espaços comerciais com mais de 5 promoções são muito penalizados (*Spot_Ratio%_Médio*=4,4% e 12% caso contrário).

Para valores de *Rat_Prog%* (T) < 12% o que influencia *Rat_Spot%* são os valores de *Rat_Ldin%* (T) e de *Tvr_Prog%* (P).

Portanto, as variáveis que mais condicionam os *ratings* comerciais são as correspondentes audiências programáticas ou adjacentes, os níveis de taxas médias de audiências para o total da TV e as durações programáticas. Excluindo estas variáveis são os factores temporais que as condicionam. Neste caso os valores de *Tempo_End* > 0,94 beneficiam *Rat_Spot%* uma vez que as telenovelas da noite pertencem a horários mais tardios e neste caso os espaços comerciais com menor número de promoções o que também está associado a uma menor duração dos espaços comerciais.

Para programas com horários de conclusão (*Tempo_End* < 0,94) antes das 22.30H os níveis de saturação publicitária influenciam os resultados e as tipologias programáticas. Uma maior saturação publicitária não penaliza *Rat_Spot%* (se *Sat_Prog* > 15%, *Rat_Spot%* *Médio* = 11,5% e 9,5% caso contrário). Para os maiores valores saem beneficiados a informação e a ficção e para os restantes valores apenas a ficção e o desporto.

Os níveis de *Tvr_Prog* (P) condicionam (se < 15%, *Spot_Ratio%* *Médio* = 1,48 e 0,84 caso contrário) a variável *Spot_Ratio%*. Os menores valores das taxas médias de audiências para o total da TV favorecem *Spot_Ratio%* uma vez que os *ratings* programáticos também são beneficiados. Os acréscimos verificados nas audiências programáticas não são

acompanhados por idênticas subidas ao nível comercial, reduzindo o valor desta variável.

Para os menores valores de *Tvr_Prog%* (P) os espaços comerciais com menos de seis promoções são os suportes mais favorecidos (*Spot_Ratio%* *Médio* = 2,41 e 1,04 caso contrário). Para os maiores valores de *Tvr_Prog%*, a duração acumulada penaliza *Spot_Ratio%* (se *Dur_Ac* > 108 seg., *Spot_Ratio%* *Médio* = 0,81 e 0,95 caso contrário). Para os menores valores de *Dur_Ac*, a duração programática parcial também condiciona os resultados (se *Dur_Prog* (P) < 380 seg., *Spot_Ratio%* *Médio* = 0,99 e 0,89 caso contrário). Portanto, as posições mais favorecidas são as iniciais (menor *Dur_Ac*) e associadas a menores durações parciais. Se os programas forem mais longos o público pode utilizar os espaços comerciais para realizar outras actividades. Para as maiores durações parciais uma vez mais *Tvr_Prog%* penaliza esta variável. Considerando as posições com maiores valores de *Dur_Ac*, *Rat_Prog%* (P) penaliza esta variável. Para valores de *ratings* parciais inferiores aos 11% os resultados diferenciam-se em função de programas em concreto.

Uma vez que os programas oscilam nos seus horários de início, os valores de *Spot_Ratio%* são tanto maiores quanto mais cedo for o seu início (se *Tempo_Time* < 0,89, 0,88; 0,75 caso contrário). Para as audiências programáticas parciais superiores a 11% as horas de início e as de conclusão condicionam os resultados.

Portanto, para este canal se concentrarmos as análises nos dois períodos de maiores audiências os níveis de *Tvr_Prog%* penalizam *Spot_Ratio%*. Para os menores valores saem favorecidos os espaços comerciais menos longos. Para os maiores valores é relevante a duração acumulada desde o início do espaço comercial até à inserção. As posições iniciais associadas a durações programáticas parciais também mais reduzidas são mais favorecidas. Para as restantes posições com maiores valores de durações acumuladas são beneficiados as inserções com menores valores de *ratings* programáticos parciais desde que o seu início seja anterior a *Tempo Time* =0,89.

V. Comparação de Resultados

As metodologias simbólicas permitem maximizar os coeficientes de determinação para ambas as variáveis dependentes consideradas, podendo ser utilizadas como metodologias de previsão se considerarmos os valores dos *ratings* comerciais. Para os valores de *Spot_Ratio%* as metodologias não lineares são as mais adequadas. No entanto as percentagens de variância explicada são reduzidas, o que nos permite confirmar os resultados de *Danaher (1995)* quanto à inexistência ou dificuldade na identificação de causas sistemáticas que permitam explicar a evolução deste indicador.

Para comparar os resultados decorrentes da aplicação da Regressão Linear Múltipla e das Árvores de Regressão CART comparamos os coeficientes de regressão estandardizados (β) com os valores da Medida M para as Árvores de Regressão. A medida M calcula-se através da seguinte expressão:

$$M(X_j) = \sum_{O \in A} z^{jO} \Delta I(\Pi^j, O, A)$$

com

$$z^{jO} = \begin{cases} 1, & \Pi^j \in \Pi^s(O) \\ 0, & \Pi^j \notin \Pi^s(O) \end{cases}$$

onde Π^j corresponde à ramificação do nó O associado com a variável explicativa X_j ; Π^s corresponde ao conjunto de variáveis substitutas e $\Delta I(\Pi^j, O, A)$ traduz o decréscimo na diversidade da Árvore global com a utilização de X_j na partição do nó O.

A medida M estandardizada corresponde a

$$M(X_j) = 100 * \left[\frac{M(X_j)}{\max M(X_j)} \right]$$

Uma das vantagens dos valores de Betas em relação à medida M utilizada nas Árvores CART traduz-se em que o seu sinal nos dá a direcção da influência da variável. A necessidade de utilizar as variáveis *dummy* na RLM dificulta a comparação dos resultados.

Verificamos que segundo as Árvores CART a variável que se associa em maior grau com os *ratings* comerciais corresponde à duração do espaço comercial e, consequentemente também ao número de inserções que os constituem.

Segue-se o dia de semana, o canal e a localização do espaço comercial (*Tabela F.7.5.11*).

CART	M(X _j)	m(X _j)
Dur_Espaço_Comercial	3	100
Num_Inserções	2,5	83
Dia	2,1	70
Pos_Ex	1,5	50
Pos_Fol	1,4	46
Loc_Espaço_Comercial	1,4	46
Canal	1,35	45
Dur_Spot	1,2	40
Dur_Ac	0,7	23
Pos_Spot	0,5	17
Ex_Promo	0,4	13
Fol_Promo	0,05	1
MLR	Beta	Beta Normalizado
Fim de Semana	-4,2	100
Desporto	-4	95
SIC	-3	71
Loc_Espaço_Comercial	-2,5	59
Sexta	-2	47
Informação	-1,6	38
Num_Inserções	-0,10	2
Dur_Spot	-0,05	1
Dur_Espaço_Comercial	-0,05	1

Tabela F.7.5.11 - Comparação de coeficientes Beta de Regressão e Medidas M do algoritmo CART - Canais privados

Na *Tabela F.7.5.12* consta a comparação da importância relativa das variáveis independentes para os canais privados segundo ambas as metodologias e na *Tabela F.7.5.13* para os canais públicos, atendendo aos dados de 2008 e para comparar os resultados.

Para os canais públicos o modelo de RLM apresenta um $R^2=0.88$ e um valor de $R^2=0.37$ para os canais privados. O algoritmo CART atinge os 85% de variância explicada para a totalidade da amostra (validação cruzada com 30% da amostra). Atendendo aos dados de 2008 verificamos que as variáveis mais relevantes para os canais privados correspondem ao número de inserções associado ao espaço comercial e à localização do espaço comercial.

Os canais privados apresentam uma audiência com um perfil mais jovem cuja tendência para o *zapping* contribui para que as audiências dos espaços comerciais entre programas sejam menores, apesar dos espaços comerciais no decorrer de um programa terem maior número de inserções e serem, consequentemente mais longos.

RLM	Beta	Beta Normalizado
Canal RTP1	5.89	100
Desporto	3.28	56
Sábado	1.97	33
Juventude	1,41	24
Quinta	1,12	19
Segunda	-1,11	19
Informação	0,83	14
Cultura	0,77	13
Ex_Promo	-0.32	5
CART	M(X _j)	m(X _j)
Canal	4.40	100
Tipologia	2.22	50
Pos_Espaço comercial	1.82	41
Pos_Ex	1.89	43
Espaço_comercial_Dur	1.79	41
Pos_Segue	1.46	33
Dia	1.26	29
Num_Inserções	1.24	28
Pos_Spot	0.46	10
Dur_Ac	0.30	7
Dur_Spot	0.13	3

Tabela F.7.5.13 - Comparação de coeficientes Beta de Regressão e Medidas M do algoritmo CART (Canais Estatais)

RLM	Beta	Beta Normalizado
Sábado	4.41	100
Diversão	4.33	98
SIC	3.45	78
Segunda	3.29	75
Pos_Espaço comercial	3.04	69
Sexta	2.19	50
Quinta	1.79	41
Informação	1.61	37
Num_Inserções	0.13	3
Dur_Spot	0.05	1
Terça	0.05	1
CART	M(X _j)	m(X _j)
Dur_Espaço comercial	3.36	100
Num_Inserções	2.64	78
Dia	2.31	69
Pos_Ex	1.88	56
Pos_Segue	1.80	54
Pos_Espaço comercial	1.79	53
Canal	1.75	52
Dur_Spot	1.31	39
Tipologia	1.08	32
Dur_Ac	0.68	24
Num_Inserções	0.46	20
Ex_Promo	0.03	14

Tabela F.7.5.12 - Comparação de coeficientes Beta de Regressão e Medidas M do algoritmo CART (Canais Privados)

As tipologias programáticas de Informação e de Ficção são as tipologias que apresentam um maior número de espaços comerciais no seu decorrer. A existência de promoções programáticas e se uma inserção sucede ou não esta tipologia de inserções publicitárias não parece influenciar os valores dos *ratings*. A diferença existente entre o número de inserções que pertence a espaços comerciais com ou sem promoções para ambas as localizações do espaço comercial não é tão relevante como sucede para os canais públicos. Os espaços comerciais entre programas apresentam um ligeiro acréscimo no número de promoções (53.54% dos inserções pertencem a espaços comerciais com promoções).

Para os espaços comerciais no decorrer de um programa, 59.14% das inserções pertencem a espaços comerciais que não têm qualquer promoção. Os espaços comerciais com um maior número de inserções apresentam menores valores de *ratings*. Para os canais privados o número de inserções no decorrer de um programa duplica o número de inserções entre programas. A duração de uma inserção publicitária também penaliza a sua taxa de audiência. Apesar das variáveis dia e tipologia ainda serem relevantes, a maior regularidade de exposição televisiva e a predisposição que se verifica para o *zapping* maximiza a importância das variáveis de planeamento. O perfil de audiências dos canais privados corresponde às classes sociais mais altas, ao sexo feminino e quanto às faixas etárias, para o canal TVI abrange dos 4 aos 14 anos ou superiores aos 35 anos e para o canal SIC abrange dos 4 aos 14 anos e as faixas etárias superiores aos 55 anos.

VI. Análises realizadas à exposição comercial dos segmentos sociodemográficos

Para a avaliação da repercussão das variáveis sociodemográficas na ordem de grandeza dos *ratings* comerciais realizaram-se análises ANOVA ou Testes T para diferenças de médias. Observam-se diferenças entre as médias observadas para os segmentos feminino e masculino (10,3% e 8,4% respectivamente). O Teste T tem um p valor de 0,0005 logo significativo e o teste de Levene também, mas com um valor elevado da dimensão dos grupos os efeitos de heterogeneidade das variâncias são mínimos.

Atendendo às Regiões Geográficas diferenciam-se dois grupos: Lisboa, Porto e Litoral Norte (8,3%) e o litoral Centro, Interior e Sul com valores superiores (10%). Como o coeficiente de correlação entre as médias e as variâncias é elevado realizou-se um teste de Kruskal-Wallis. Às cidades metrópoles, costumam estar associados menores valores de *ratings* comerciais, mas, também, de programas.

Quanto mais reduzida for a classe social maiores são os valores de *ratings*, não se observando diferenças estatisticamente significativas entre as classes C1 e a C2. A classe AB que corresponde à mais elevada apresenta uma menor exposição televisiva, e, logo, também comercial.

As pessoas com um maior rendimento veem menos televisão. Em princípio quando assistem é porque esperam algum tipo de gratificação inerente à programação, consequentemente encaram os espaços comerciais de forma negativa. Além disso quem é mais exigente relativamente à publicidade é também mais susceptível a ficar decepcionado quando ela não corresponde às suas expectativas.

Para os telespectadores que não gostam do formato ou frequência da publicidade, a sua intensidade será tanto maior quanto maior for a classe social. Quem não gosta da programação, também não gosta dos conteúdos transmitidos, encara a publicidade da mesma forma. No entanto quem aprecia a programação poderá ver os espaços comerciais como repetitivos, apresentando imagens irreais, interferindo com a programação ou não lhes oferecendo qualquer tipo de benefício.

Em relação às faixas etárias diferenciam-se os seguintes segmentos: 15-34 com os menores valores, segue-se 4-14 com 35-54; 55-64 e +64. Com o aumento da idade também aumenta a média associada às audiências comerciais. As reduzidas audiências verificadas para o segmento dos 15 aos 34 anos são devidas quer a uma maior predisposição para realizar o *zapping*, como também por ser menor o seu consumo televisivo, mesmo ao nível do prime time.

Segundo *Burnett (1991)*, os mais idosos preferem obter por si só informações relativas aos produtos que pretendem adquirir, não atribuindo grande relevância aos espaços comerciais. Esta constatação é atribuída à sua incapacidade de processar a informação rapidamente; à dificuldade em recordar a informação e à tendência para serem facilmente convencidos e ludibriados.

A sua atitude face à publicidade é negativa, consequentemente o facto de lhes estar associado elevados valores de ratings comerciais não significa que na realidade as suas decisões de compra sejam influenciadas pela publicidade.

Esta visão negativa, de que a publicidade é inútil e enganosa é preponderante no sexo masculino. Outro factor que favorece uma visão negativa da publicidade é o fenómeno designado de “Advertising Clutter”, ou seja, a proliferação de um elevado número de anúncios de curta duração. Afecta sobretudo os mais idosos e os indivíduos cognitivamente mais lentos. O “clutter” é um fenómeno complexo, que se encontra associado quer com a duração das inserções, como também com a duração dos espaços comerciais.

Segundo *Johnson e Cobb-Walgren (1994)*, o aumento da idade diminui a rapidez cognitiva, afectando a taxa de recordação, quanto maior for o “clutter” dos dados.

Algumas soluções encontram-se sugeridas por Ray e Webb (1986), como, por exemplo utilizar inserções que estimulem as pessoas a pensar sobre os produtos e a respeito do seu impacto nas suas vidas; repetir os anúncios frequentemente; utilizar inserções um pouco mais longas (30 seg. em vez dos 15 seg.); colocar inserções em programas cujo padrão de intervalos seja variado, aos quais estejam associados espaços comerciais de curta duração e de preferência em programas semanais e em eventos especiais.

De acordo com Danaher (1995) as faixas etárias mais jovens evitam os espaços comerciais em maior grau, mas os homens não apresentam maior tendência para o zapping. Segundo o autor, na realidade os padrões de zapping não são explicados com base em características pessoais e do lar, não existem padrões típicos de *zapping*.

Os telespectadores evitam os espaços comerciais quando lhes apetece, mas a maioria não os evita mudando de canal. Também verificou que as áreas rurais apresentam maior exposição face às metropolitanas, assim como uma correspondência inversa entre o nível de rendimento e os valores dos *ratings*.

“... Since the overwhelming majority of ad avoidance behaviour was unexplained by personal and household demographics, no typical pattern of ad avoiders emerges. The fact that viewers avoid ads when they feel like it, but the vast majority of the time they do not avoid television ads by switching channels.” Danaher (1995).

Com base nos dados nacionais verificamos que atendendo à elevada correlação entre ratings comerciais e programáticos, a diferenciação dos níveis de *ratings* comerciais pode ser atribuída às diferenças entre a exposição ao meio televisivo e programático.

APÊNDICE G

ÁREAS METODOLÓGICAS

G.1. INTRODUÇÃO

Verificamos que existe uma grande diversidade de abordagens analíticas que podem contribuir para o desenvolvimento das análises de previsão e explicativas das audiências televisivas, como, por exemplo, as que constam do seguinte *Quadro G.1.1*.

As Técnicas Estatísticas que expressam modelos mediante fórmulas e funções, por exemplo, a Regressão linear global ou local e a simbólica.

As técnicas Bayesianas que estimam a probabilidade de pertença mediante a estimação de probabilidades condicionadas recorrendo ao Teorema de Bayes, por exemplo, as Redes Bayesianas.

As técnicas de Regras de Associação baseadas em frequências que contam o número de vezes que duas ou mais observações ocorrem conjuntamente, por exemplo, o algoritmo Apriori.

As técnicas de Redes Neurais que aprendem um modelo em que a tipologia da Rede e o peso das conexões determina qual o padrão apreendido.

As técnicas de *Clustering* baseadas em observações, na densidade ou distância, por exemplo, K Média, KNN.

As técnicas baseadas em Árvores e Regras de Decisão com uma representação na forma de Regras de classificação ou de previsão.

As técnicas Relacionais, sejam estas Árvores ou outros algoritmos de classificação que representam modelos como linguagens lógicas, funcionais e lógico-funcionais.

Quadro G.1. Áreas metodológicas para o planeamento de *media*.

Fonte: Elaborado pela autora.

As técnicas de modelização estatística paramétricas são eficientes e em muitos casos compreensíveis, enquanto que as técnicas não paramétricas são mais expressivas. Ambas trabalham facilmente com dados numéricos.

As técnicas Bayesianas são fáceis de usar, eficientes, lidam com muitos atributos, apresentam robustez ao ruído, a sua expressividade é limitada e dependem da discretização. Estas abordagens não constroem modelos, estimam uma série de probabilidades, com excepção da Rede Bayesiana que poderá ser muito expressiva.

As Árvores de Decisão e as Regras são fáceis de usar, admitem todo o tipo de variáveis, lidam bem com o ruído e com os valores omissos. Também são eficientes e inteligíveis, podendo-se expressar como um conjunto de Regras. Apresentam uma limitada expressividade e são instáveis perante variações na amostra. As técnicas relacionais são expressivas e permitem o tratamento de dados com estruturas e captam padrões relacionais e recursivos, assim como expressam conhecimento prévio sob a forma de regras. Apresentam alguma dificuldade na sua manipulação e pouca eficiência.

As Redes Neurais na sua construção requerem alguma experiência por parte do seu utilizador para lhes conseguir retirar o maior partido. Quando estão bem ajustadas obtêm-se resultados com uma elevada precisão, são expressivas e capturam modelos não lineares. Porém, apresentam sensibilidade a valores anómalos, precisam de uma elevada dimensão da amostra, são abordagens lentas e até certo ponto incompreensíveis.

As técnicas baseadas na distância são fáceis de usar, eficientes se o número de observações não for elevado, apresentam bastante expressividade, mas não constroem um modelo, por conseguinte não produzem conhecimento compreensível.

As previsões finais na área em análise devem ser realizadas pelos utilizadores incorporando a sua experiência, intuição e conhecimento sobre o *media*. Os modelos consequentemente devem ser claros e simples, constituindo um ponto de partida, sobre os quais realizar alguns ajustes finais e que são função dos contextos.

Para serem realizadas alterações nos resultados proporcionados será necessário que os seus utilizadores consigam perceber como os modelos funcionam e que cálculos realizam. Caso contrário não será possível avaliar se os resultados são erróneos ou coerentes. Verificámos nas análises realizadas que as metodologias de *Data Mining* produzem resultados satisfatórios atendendo à dimensão dos erros que constam da bibliografia da especialidade.

Nos seguintes sub-capítulos apresentam-se algumas considerações a respeito de algumas abordagens metodológicas implementadas nesta tese.

G.1.1. Comparação dos algoritmos CHAID e CART de Árvores de Regressão

Ambos os algoritmos CHAID e CART permitem o desenvolvimento de Árvores de Regressão, contudo apresentam algumas diferenças.

Para evitar o sobre ajustamento aos dados, ambas as metodologias restringem a dimensão da Árvore final durante a fase de treino. Enquanto, o algoritmo CHAID recorre a uma regra estatística de paragem inibindo o total desenvolvimento da Árvore, o algoritmo CART utiliza a poda. A poda da Árvore é realizada avaliando o seu desempenho numa amostra de teste e comparando-o com os resultados alcançados na amostra de treino. Neste processo alguns ramos da Árvore são eliminados até ambos os desempenhos serem equivalentes.

O CHAID recorre a uma única amostra para alcançar a solução, enquanto o CART recorre à amostra de treino para a construção do modelo e a uma amostra de teste para realizar a poda. O CHAID em cada *split* procura a variável que se seleccionada explica em maior grau a variável critério. Recorre a testes de hipóteses de independência entre a variável explicativa e a variável critério, nomeadamente ao Teste do Qui-Quadrado ou ao Teste F. Quando as variáveis são independentes, o algoritmo para o desenvolvimento da Árvore.

O CART por sua vez decide quais são os *splits* a realizar com base no grau de redução de uma medida de diversidade. Posteriormente, os *splits* são reavaliados em função do grau de ajustamento aos dados. Consequentemente, o CHAID é mais relevante para a análise das bases de dados e o CART para promover a sua previsão.

O CHAID é mais adequado quando o objectivo corresponde à descrição e caracterização das relações existentes entre as variáveis explicativas e a variável critério, enquanto o CART se deve utilizar para a elaboração de um modelo que permita uma correcta classificação ou previsão de novas observações. O CART apenas permite *splits* binários, enquanto o CHAID aceita *splits* múltiplos, sendo este último algoritmo frequentemente utilizado na área do *Marketing*, nomeadamente quando se pretende a segmentação dos dados, enquanto nas restantes áreas o CART em geral é o algoritmo mais utilizado.

Em análises de regressão o CHAID, recorrendo a *splits* não binários conduz a Árvores mais dispersas. Permite que inúmeros nós terminais estejam conectados a um único ramo, o que poderá ser facilmente resumido em quadros que cruzam as múltiplas categorias de cada uma das variáveis envolvidas nas análises.

O CART construindo Árvores binárias não pode ser resumido de forma tão eficiente dificultando a interpretação e apresentação das conclusões que é possível retirar. Quando o objectivo das análises é essencialmente explicativo, sendo mais relevante as interacções observadas entre as variáveis o CHAID corresponde à metodologia mais adequada. O CART tende a seleccionar variáveis que são utilizadas inúmeras vezes em diferentes fases do desenvolvimento da Árvore, consequentemente as conclusões nem sempre são claras, coerentes e fiáveis para a previsão ou classificação de novas observações. O CART corresponde a um algoritmo mais complexo e a sua execução poderá ser demorada quando se lida com um grande número de observações.

Uma das limitações do CHAID corresponde a que os seus segmentos se baseiam em uma única variável critério. Em contextos multicritério, não fica claro como obter uma segmentação global dos dados. Em função da variável seleccionada, assim, podem variar os segmentos finais obtidos por aplicação desta metodologia.

Além do mais, as categorias das variáveis explicativas podem ser agregadas de forma distinta consoante a variável critério utilizada, originando diferentes segmentos. Quando existem inúmeras variáveis dependentes, elas podem ser de diferentes tipologias. Os segmentos resultantes da análise de variáveis ordinais podem ser distintos face aos que resultariam de serem consideradas variáveis do tipo nominal.

Quando está em causa a previsão de dados será preferível recorrer a diferentes algoritmos, comparar os seus resultados e, assim, decidir sobre quais são os resultados mais razoáveis e que apresentam um melhor desempenho por análise dos erros de previsão.

G.1.2. Comparação de Árvores de Regressão e a Regressão linear Múltipla

Na Regressão Linear, o modo *stepwise* de selecção das variáveis tende a conduzir a um sobre ajustamento aos dados. Sistemáticamente, tende a incluir demasiadas variáveis explicativas e sobrestima a proporção de variância explicada. O algoritmo CART evita o sobre ajustamento aos dados através da poda que se traduz na eliminação de ramos de modo a reduzir a dimensão da Árvore e, assim, obter resultados generalizáveis a outras amostras, com perda mínima de capacidade explicativa.

Uma vez obtida a Árvore com a maior variância explicada, recorrendo à poda evita-se a perda de informação decorrente de uma paragem demasiado breve do seu crescimento ultrapassando o problema associado aos algoritmos pré-poda que estão presentes, por exemplo, no algoritmo CHAID. Nas Árvores de Regressão não é necessário a transformação inicial das variáveis, sejam estas métricas (para se verificarem os pressupostos paramétricos no caso da Regressão é necessário a sua transformação) ou nominais (codificação com variáveis *dummy* na aplicação da Regressão Linear).

As Árvores são robustas à presença de observações atípicas, uma vez que os *splits* são realizados para valores não extremos, e estas observações são conduzidas para os nós extremos não influenciando a Árvore remanescente, enquanto na Regressão Linear é necessário proceder à sua exclusão.

As Árvores permitem o tratamento de bases de dados de grande dimensão. Para a aplicação do algoritmo CHAID será necessário uma base de dados com um elevado número de observações, uma vez que a utilização de *splits* múltiplos pode levar à sua fragmentação antes de se detectar a estrutura subjacente.

A este nível o CART sendo binário detecta a estrutura existente antes de se dispor de poucas observações, no entanto tende a utilizar a variável que permite um maior número de *splits* no processo, conduzindo a resultados por vezes pouco fiáveis e claros.

Para a Regressão a dimensão amostral influencia a potência dos testes, consequentemente uma elevada dimensão da amostra poderá ser problemática, pois os testes tornam-se sensíveis e podem indicar a significância de qualquer relação observada por menor que esta seja.

Os modelos paramétricos de regressão não têm em consideração as interacções existentes entre as variáveis explicativas. Os termos de interacção entre duas ou mais variáveis podem ser incluídos, no entanto a metodologia não ajuda na sua identificação, pois apenas revela uma estrutura dominante. A complexidade dos cálculos e a dificuldade de interpretação dos termos de maior ordem não torna atractiva a sua utilização. As Árvores recorrendo simultaneamente a variáveis contínuas e nominais são mais eficazes, quer na identificação de variáveis explicativas chave como também na identificação das respectivas interacções.

O algoritmo CART sendo binário tende a utilizar a mesma variável em vários níveis o que poderá conduzir a resultados pouco claros. No entanto, este problema pode em parte ser contornado, recorrendo ao conceito de surrogate (inexistente no CHAID) segundo o qual é possível o cálculo de uma medida de importância relativa global das variáveis explicativas, ou seja, o cálculo de uma medida equivalente aos Betas da Regressão Linear Múltipla.

O conceito de *surrogate* traduz-se em *splits* alternativos ao *split* primário no caso de uma observação apresentar um valor omisso para a variável que conduz a uma maior redução da medida de diversidade, ou seja, para o *splitter* primário.

Se os *splits* forem univariados (também se podem utilizar combinações lineares das variáveis métricas nas partições dos dados), as variáveis podem ser ordenadas numa escala de 0 a 100, em termos da sua potencial capacidade explicativa. Comparando as duas medidas, ou seja, os Betas e a medida de *Breiman et al. (1984)* existem diferenças obvias. Os modelos de Regressão, sendo lineares procuram determinar uma fórmula quantitativa que descreva a relação existente entre a variável dependente e as explicativas, determinando a magnitude, extensão e direcção (dado pelo sinal dos Betas de cada variável) da relação existente.

Os coeficientes betas quantificam a variação observada na variável dependente devido a uma variação unitária verificada na variável independente. A medida de *Breiman et. al (1984)* caracteriza a repercussão das variáveis explicativas como uma medida do grau em que estas permitem uma redução da diversidade e risco associado à previsão.

A previsão nas Árvores é rápida, não sendo necessário realizar grandes cálculos. Nas Árvores a previsão associada a uma observação será o valor médio da variável dependente para o nó em que essa observação se encontra incluída. Todas as observações presentes num nó obedecem à mesma regra proposicional, ou seja, à combinação lógica de valores dos atributos. No caso da Regressão a previsão passa por um único valor contínuo. O conceito de *surrogate*, ou seja, de variáveis substitutas, criando *splits* alternativos permite o tratamento de valores omissos e pode ajudar na identificação de mecanismos menos onerosos de previsão em contextos de maiores custos associados ao *splitter* primário. Perante a ausência de determinada informação, ainda se poderá realizar previsões ponderando os nós da subárvore que for possível atingir.

Quando se aplica a Regressão, se surgirem situações de valores omissos, será necessário a sua eliminação ou recorrer a um método de cálculo, como, por exemplo, a sua substituição pela média, um valor constante, entre outras. Os resultados que decorrem da aplicação das Árvores são geralmente mais lógicos e consistentes face aos das metodologias paramétricas.

As Árvores correspondem a uma metodologia relativamente automática onde não é necessário a especificação de grande *input* por parte do analista e de fácil interpretação por parte de não estatísticos. Contrastando com outras metodologias multivariadas que requerem um grande esforço na análise dos resultados e na especificação do modelo.

Uma das desvantagens da Regressão Linear traduz-se no facto de corresponder a um modelo global, onde uma única função de previsão abrange todo o domínio. Quando os dados apresentam inúmeras tendências, que interagem com complexidade e não linearidade a sua síntese quando viável poderá ser extremamente difícil e confusa.

Uma alternativa será a aplicação da Regressão não linear que consiste na subdivisão ou partição do espaço de entrada em regiões de menor dimensão, onde as interacções sejam mais facilmente manejáveis. Subdivide-se novamente esses espaços – a denominada partição recursiva - até atingirmos um conjunto de espaços onde seja possível ajustar modelos mais simples. O modelo global apresenta duas componentes: a partição recursiva e o ajustamento de um modelo mais simples a cada célula resultante da partição. As Árvores de Regressão representam a partição recursiva.

G.1.3. Comparação das Redes Neurais RBF e MLP

Ambas as Redes Neurais RBF e MLP são do tipo supervisionado e consideradas como aproximadores universais. As neuronas ocultas nas Redes MLP dependem de somas ponderadas dos valores de entrada, transformados por funções de activação monótonas. Uma função de activação geralmente aplicada é a função sigmoideal que é não-linear e continuamente diferenciável, como, por exemplo, a função Logística e a função Tangente Hiperbólica. Nas Redes RBF a activação de uma unidade oculta é determinada por uma função não linear da distância entre o vector de entrada e um vector de referência. As unidades ocultas de uma Rede RBF apresentam funções de activação que são localizadas e apresentam uma base radial sobre o seu domínio. Uma Rede MLP cria uma representação distribuída no espaço de valores de activação para as unidades ocultas, pois para um vector de entrada, muitas unidades ocultas contribuirão para a determinação do seu valor de saída.

A “interferência” e o “acoplamento cruzado” entre as unidades escondidas levam a resultados que são não-lineares, resultando em problemas de mínimos locais ou em regiões quase planas na função de erro, factores que podem levar a uma convergência muito lenta do algoritmo. As Redes RBF apresentam funções de base localizadas que formam uma representação no espaço de unidades escondidas que é local em relação ao espaço de entrada porque para um vector de entrada, tipicamente apenas algumas unidades escondidas apresentarão activações significantes. Consequentemente, as Redes RBF tendem a produzir aproximações locais.

As Redes MLP apresentam inúmeras camadas de pesos e um complexo padrão de conectividade, de modo que nem todos os possíveis factores de ponderação em determinada camada podem estar presentes. Podem ser utilizadas diferentes funções de activação na mesma Rede. Uma Rede RBF apresenta uma arquitectura simples, estando constituída por duas camadas de pesos, em que a primeira contém os parâmetros das funções de base radial e a segunda calcula combinações lineares das activações das funções de base radial para gerar os valores de saída.

Os parâmetros de uma Rede MLP são usualmente determinados ao mesmo tempo, correspondendo a uma estratégia global de treino de carácter supervisionado. Este tipo treino apresenta um elevado custo computacional, pela necessidade de retropropagação do erro, o que faz com que estas Redes apresentem uma aprendizagem muito lenta. Porém o seu desempenho de generalização, é bom. Uma Rede RBF é treinada em duas etapas, com as funções de base radial, sendo determinadas em primeiro lugar por técnicas não-supervisionadas, usando para tal os dados de entrada. Na segunda camada, que abrange os factores de ponderação, é determinada por métodos lineares supervisionados e de rápida convergência. A diferente estratégia de treino e a consequente diferença de velocidade de treino entre as duas tipologias de redes faz com que as MLP sejam menos adequadas do que as redes RBFs, em operações dinâmicas que envolvam uma aprendizagem continuada, como, por exemplo, para a previsão das séries temporais e para as aplicações on-line.

Num contexto de aproximação de funções, sob idênticas condições do ambiente no qual estão inseridas, de uma forma geral pode-se afirmar que o erro final alcançado por uma Rede RBF é menor face ao de uma Rede MLP; a convergência de uma RBF pode chegar a uma ordem de grandeza mais rápida do que a convergência de uma Rede MLP; a capacidade de generalização da rede MLP é em geral superior á capacidade de generalização da RBF. Uma das vantagens das Redes RBF em relação às Redes MLP traduz-se no seu tempo de treino mais reduzido.

Contudo, as Redes Neurais que utilizam a retro-popagação dos erros, assim como muitas outras tipologias de Redes neuronais, correspondem a "caixas pretas", nas quais quase que se desconhece porque a Rede chega a determinado resultado, uma vez que os modelos não apresentam justificativas para as suas soluções. Neste sentido, esforços têm sido desenvolvidos com o intuito de viabilizar a extracção de conhecimento das Redes Neurais Artificiais, e na criação de procedimentos explicativos, onde se tenta justificar o comportamento da Rede perante determinadas situações.

Uma outra limitação refere-se ao tempo de treino das Redes recorrendo à retropopagação dos erros que tende a ser demasiado lento. Algumas vezes são necessários milhares de ciclos para se atingir níveis de erros aceitáveis, principalmente sendo simulado em computadores seriais, pois a CPU deve calcular as funções para cada unidade e para as suas conexões separadamente, o que poderá ser problemático em Redes de elevada dimensão, ou quando se tem que processar uma grande quantidade de dados. Será muito difícil definir a arquitectura ideal da Rede de forma que ela seja tão grande quanto o necessário para conseguir obter as representações necessárias, mas simultaneamente pequena o suficiente para se alcançar um menor tempo de treino. Não existem regras claras para se definir quantas unidades devem existir nas camadas intermédias ou quantas devem ser as camadas, ou como devem ser as conexões entre essas unidades. Para resolver estas questões, algoritmos Genéticos podem ser utilizados para encontrar automaticamente boas arquitecturas de Redes Neurais, eliminando muitas armadilhas associadas às abordagens da Engenharia humana

