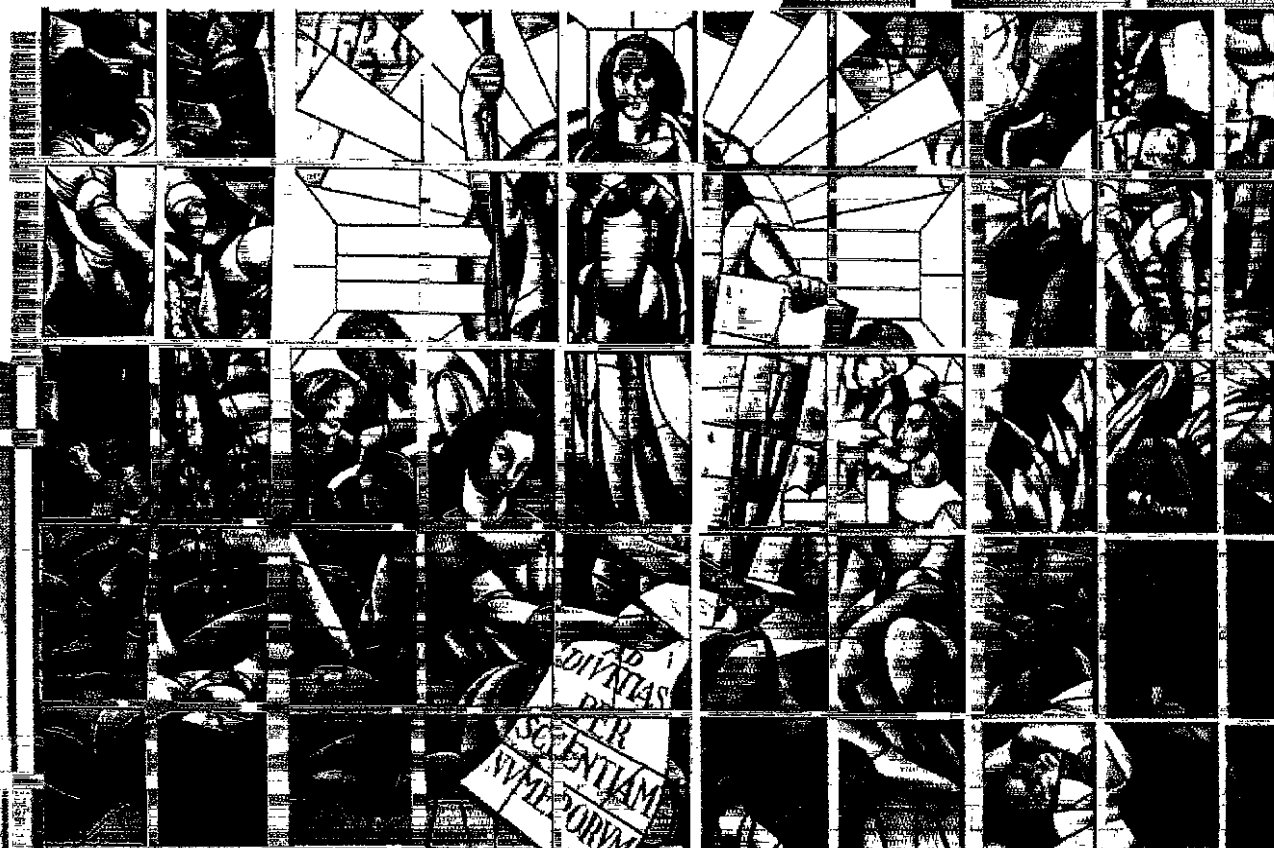




INSTITUTO NACIONAL DE ESTATÍSTICA

PORTUGAL

REVISTA DE ESTATÍSTICA



VOLUME III

3º QUADRIMESTRE 1999

CATALOGAÇÃO RECOMENDADA:

REVISTA DE ESTATÍSTICA. Lisboa, 1996-
Revista de estatística / ed. Instituto Nacional de
Estatística. - Vol. 1, 1º quad. 1996-
Lisboa : I.N.E., 1996- . - 30 cm
Quadrimestral
ISSN 0873-4275

FICHA TÉCNICA

- DIRECTOR
 - *Adrião Simões Ferreira da Cunha*
- DIRECTOR-ADJUNTO
 - *Pedro Jorge Nunes da Silva Dias*
- CONSELHO EDITORIAL
 - *Adrião Simões Ferreira da Cunha*
 - *António Daniel Correia dos Santos*
 - *Dinis Duarte Ferreira Pestana*
 - *Francisco José Neto Melro*
 - *João António Branco*
 - *João Ferreira do Amaral*
 - *Oscar Soares Barata*
 - *Pedro Jorge Nunes da Silva Dias*
 - *Pedro Miguel Girão Nogueira Ramos*
 - *Sérgio Manuel Bacelar e Silva*
- SECRETARIADO DE REDACÇÃO
 - *Eduarda Lília Marques Martins*
- EDITOR
 - *Instituto Nacional de Estatística*
Av. António José de Almeida, n.º 2
1 000-043 LISBOA
Tel.: (01) 842 61 00
Fax: (01) 842 63 64
- CAPA
 - *Design de Mário Bouçados sobre o vitral do INE da autoria do pintor Abel Manta*
- LAYOUT E MAQUETAGEM
 - *Mário Bouçados*
- IMPRESSÃO
 - *Instituto Nacional de Estatística*
Secção de Artes Gráficas
- TIRAGEM
 - *750 exemplares*
- DEPÓSITO LEGAL
 - *N.º 99514/96*

PREÇO (IVA 5% incluído)

- N.º avulso 2 500\$00
- Assinatura annual 6 000\$00

© INE, Lisboa, Portugal, 1999* Reprodução autorizada, excepto para fins comerciais, com indicação da fonte bibliográfica

CREDITS

- DIRECTOR
 - *Adrião Simões Ferreira da Cunha*
- ASSISTANT DIRECTOR
 - *Pedro Jorge Nunes da Silva Dias*
- EDITORIAL BOARD
 - *Adrião Simões Ferreira da Cunha*
 - *António Daniel Correia dos Santos*
 - *Dinis Duarte Ferreira Pestana*
 - *Francisco José Neto Melro*
 - *João António Branco*
 - *João Ferreira do Amaral*
 - *Oscar Soares Barata*
 - *Pedro Jorge Nunes da Silva Dias*
 - *Pedro Miguel Girão Nogueira Ramos*
 - *Sérgio Manuel Bacelar e Silva*
- EDITORIAL SECRETARIAT
 - *Eduarda Lília Marques Martins*
- PUBLISHER
 - *Instituto Nacional de Estatística*
Av. António José de Almeida, n.º 2
1 000-043 LISBOA
Tel.: (01) 842 61 00
Fax: (01) 842 63 64
- COVER DESIGN
 - *Designed by Mário Bouçados on the stain glass window at INE by the painter Abel Manta*
- LAYOUT AND GRAPHIC DESIGN
 - *Mário Bouçados*
- PRINTING
 - *Instituto Nacional de Estatística*
Secção de Artes Gráficas
- EDITION
 - *750 exemplares*
- LEGAL DEPOSIT REGISTRATION
 - *N.º 99514/96*

PRICE (IVA 5% included)

- Single issue PTE 2,500
- Annual subscription PTE 6,000

© INE, Lisbon, Portugal, 1999* Reproduction authorised except for commercial purposes by indicating the source.



REVISTA DE ESTADÍSTICA

ÍNDICE

INDEX

- ARTIGOS
ARTICLES:

INSUCESSO NO ENSINO-APRENDIZAGEM DA MATEMÁTICA - EFEITO NAS OPÇÕES VOCACIONAIS DOS ALUNOS	
<i>Teaching-Learning Unsuccess In Mathematics Effect in Students' Vocational Options</i>	
Por/By: PAULO GOMES, INÉS QUEIRÓS.....	5
APPLICATION OF THE MODIFIED "PEAK" TO CHARACTERIZE THE M G ∞ QUEUE BUSY CYCLE LENGTH DISTRIBUTION	
APLICAÇÃO DO "PICO" MODIFICADO NA CARACTERIZAÇÃO DA DISTRIBUIÇÃO DO COMPRIMENTO DO CICLO DE OCUPAÇÃO DA FILA DE ESPERA M G ∞	
Por/By: MANUEL ALBERTO MARTINS FERREIRA.....	47
FACTORES EXPLICATIVOS DO CONSUMO DE TABACO NUM DISTRITO DO INTERIOR DE PORTUGAL	
DETERMINANTS OF THE CONSUMPTION OF TOBACCO IN AN INTERIOR DISTRICT OF PORTUGAL	
Por/By: JOSÉ RAMOS PIRES MANSO.....	57
NADOS VIVOS: ANÁLISE E ESTIMAÇÃO	
LIVE BIRTHS: ANALYSIS AND ESTIMATION	
Por/By: TERESA BAGO D' UVA.....	79

INFORMAÇÕES

- *INFORMATIONS:*

ACTIVIDADES E PROJECTOS IMPORTANTES NO ÂMBITO DO SISTEMA ESTATÍSTICO NACIONAL	
<i>IMPORTANT ACTIVITIES AND PROJECTS IN THE SCOPE OF THE NATIONAL STATISTICAL SYSTEM.</i>	117
CONGRESSOS, SEMINÁRIOS, COLÓQUIOS E CONFERÊNCIAS	
<i>CONGRESS, SEMINARS AND CONFERENCES.</i>	123
ACÇÕES DESENVOLVIDAS PELO INE NO ÂMBITO DA COOPERAÇÃO BILATERAL E MULTILATERAL	
<i>ACTIONS ACHIEVED BY NSI IN THE SCOPE OF BILATERAL AND MULTILATERAL COOPERATION.</i>	135
FUNDAMENTO, OBJECTO E ÂMBITO DA REVISTA.	
<i>FOUNDATION, SUBJECT MATTER AND SCOPE OF THE REVIEW.</i>	141
NORMAS DE APRESENTAÇÃO DE MANUSCRITOS PARA A REVISTA. . . .	
<i>RULES FOR SUBMITTING MANUSCRIPTS TO THE REVIEW.</i>	143

INSUCESSO NO ENSINO-APRENDIZAGEM DA MATEMÁTICA - EFEITO NAS OPÇÕES VOCACIONAIS DOS ALUNOS

Autores:
Paulo Gomes
e
Inês Queirós

INSUCESSO NO ENSINO-APRENDIZAGEM DA MATEMÁTICA EFEITO NAS OPÇÕES VOCACIONAIS DOS ALUNOS

TEACHING-LEARNING UNSUCCESS IN MATHEMATICS EFFECT IN STUDENTS' VOCATIONAL OPTIONS

Autores: Paulo Gomes

- Director Regional do Instituto Nacional de Estatística – DRN.

e

Inês Queirós

- Professora do Quadro de Nomeação Definitiva da Escola Secundária de Augusto Gomes – Matosinhos.

- Assistente do 1º Triénio do Instituto Superior Politécnico Portucalense.

RESUMO:

- A par do sentimento geral de que a formação em Matemática é básica para outras disciplinas, sabe-se que as opções vocacionais dos alunos são em muitos casos determinadas pelo rendimento obtido em Matemática durante o seu percurso nos primeiros anos de escolaridade. Conscientes da gravidade e da dimensão deste problema, procuramos com este estudo dar alguma contribuição para a sua caracterização, acreditando na possibilidade de superação destes fenómenos de enviesamento social.

As razões apontadas pelos alunos para os seus fracos desempenhos em Matemática são de natureza diversa. Este estudo toma como ponto de partida a aplicação de um questionário a uma amostra representativa de alunos do 10º ano de escolaridade das Escolas Públicas do Distrito do Porto. Pretende-se fundamentalmente analisar as associações entre um conjunto de variáveis eventualmente relacionadas com o insucesso escolar dos alunos em Matemática, visando assim conhecer os contornos mais gerais daquele problema.

PALAVRAS-CHAVE:

- *Insucesso escolar em Matemática, opção vocacional, análise estatística univariada, análise estatística bivariada e multivariada.*

ABSTRACT:

- It's a general feeling that learning Mathematics is fundamental to other scientific areas. It's known that students' vocational option is determined by their own progress during the first years of studying Mathematics. Being aware of the seriousness and the extension of this problem we try with this work to give some contribution to its better definition. We believe it's possible to help students choose their own way facing Mathematics, and surpass these kind of social twist phenomena.

The reasons mentioned by students, for their poor performance in Mathematics are due to several reasons. This study started with a quiz test applied to a representative sample of 10th form students in state highschools in the district of Oporto. We aim to examine the connection among several variables, eventually related to students' scholar unsucess in Mathematics. It is possible, at the end of this work, to know better the outline of this particular problem.

KEY-WORDS:

- *Mathematics scholar insucess, vocational option, univariate statistical analysis, bivariate and multivariate statistical analysis.*

1. INTRODUÇÃO

O desempenho em Matemática tem constituído um critério decisivo para seleccionar os alunos, especialmente no que se refere ao acesso às profissões de natureza técnica e científica. Aqueles que obtêm maus resultados nesta disciplina desencorajam-se de enveredar por uma carreira de engenharia ou por um curso de ciências. Implicitamente, a Matemática torna-se determinante nas opções profissionais de muitos alunos.

A convicção generalizada sobre as dificuldades e deficiências que os alunos demonstram no domínio da Matemática e as consequências inerentes a tal situação, tem cada vez mais preocupado a sociedade, nomeadamente aqueles que, conscientes da importância que esta disciplina constitui no desenvolvimento formativo, cultural, social e político do aluno, procuram estudar estratégias de melhoria da qualidade de ensino-aprendizagem da Matemática, empenhando-se assim na sua renovação.

Neste artigo o insucesso escolar será entendido no sentido mais geral, ou seja, como sinónimo de não aprovação. Será tratado o insucesso na disciplina de Matemática, procurando-se analisar as causas que os alunos atribuem a tal situação.

Na secção 2 descrevemos todo o trabalho de campo realizado com base num questionário aplicado a uma amostra de alunos que frequentaram o 10^o ano de escolaridade em Escolas Públicas do Distrito do Porto, no ano lectivo de 1996/1997. Com este questionário pretendemos essencialmente conhecer a opinião dos alunos em relação ao processo de ensino-aprendizagem da Matemática, nomeadamente no que se refere ao relacionamento entre professor e aluno, capacidades científica e pedagógica do professor avaliadas pelo aluno e importância das metodologias e estratégias usadas na sala de aula. Foi também nosso objectivo analisar o efeito do rendimento dos alunos sob as suas opções vocacionais no ensino secundário. Procuramos ainda estudar a associação entre o nível sócio-profissional familiar do aluno e o seu desempenho em Matemática.

Antes do tratamento dos dados relativos ao questionário, descrevemos com algum detalhe as etapas principais deste estudo até à aplicação do questionário, como sendo a definição dos objectivos do mesmo, a determinação da base de sondagem, a concepção e redacção do questionário, a sua pré-testagem e ajustamento, a definição da amostra-piloto e por fim a determinação da dimensão da amostra final. Houve naturalmente a necessidade de tomarmos decisões quanto ao processo de amostragem mais adequado, já que a inquirição de todas os indivíduos da população em causa seria impraticável, não só por ser um processo muito moroso e dispendioso mas também porque acreditamos que as técnicas de inferência estatística hoje conhecidas nos permitem ultrapassar com grande confiança estas limitações. Assim, depois de conhecermos alguns aspectos gerais que pensamos de algum modo caracterizar a população em relação às variáveis em estudo, optamos por uma amostra estratificada, pois tivemos a convicção de que seria este tipo de amostra que melhor representaria a população sobre a qual incide o nosso estudo.

Depois dos dados recolhidos e organizados, a fase seguinte foi a construção de uma base de dados onde pudéssemos reunir toda a informação retirada dos questionários. Com o apoio de *software* estatístico adequado iniciamos a análise dos dados, começando por uma abordagem descritiva de cada uma das variáveis em estudo (secção 3), seguindo-se o estudo bidimensional (secção 4), desde a construção de tabelas de frequências, aos testes de hipóteses não paramétricos associados, uma vez que as variáveis em estudo são, na sua maioria, do tipo nominal ou ordinal.

Estudadas as associações entre pares de variáveis em estudo, passamos à análise multivariada dos dados (secção 5), pois pretendemos ainda estudar as relações entre múltiplas variáveis qualitativas, assim como as relações entre indivíduos caracterizados por duas ou mais variáveis. Sendo o objectivo final da estatística multivariada a simplificação dos dados, isto é, “sumariar um conjunto vasto de dados através de um número relativamente reduzido de parâmetros” (Reis, 1997, p. 28), procuraremos representar graficamente os indivíduos, as variáveis ou as modalidades de variáveis e descrever o conjunto inicial a partir de um menor número de “novas variáveis”, sem perda significativa da informação contida neste conjunto. A leitura destes gráficos permite-nos estabelecer “semelhanças” ou “contrastos” entre os indivíduos observados e pôr em evidência “ligações entre as variáveis” (Gomes, 1993, p. 13). Faremos uma análise exploratória com o objectivo de reduzir a dimensão dos dados, recorrendo para o efeito à análise factorial das correspondências múltiplas (AFCM), sendo este o método privilegiado de estatística descritiva multidimensional para o tratamento de inquéritos, envolvendo variáveis qualitativas ou quantitativas. Admitindo outras abordagens, nomeadamente como um caso particular da análise discriminante ou da análise canónica, a análise factorial das correspondências utilizada neste estudo será considerada como um caso particular da análise em componentes principais (ACP), mediante uma transformação preliminar do quadro inicial e uma escolha adequada das métricas Q e D no “espaço dos indivíduos” e no “espaço das variáveis”, respectivamente. Terminamos com a apresentação das conclusões finais retiradas deste estudo.

2. INSUCESSO NO ENSINO-APRENDIZAGEM DA MATEMÁTICA TRABALHO DE CAMPO

2.1 OBJECTIVOS DO ESTUDO

Esta investigação procura estudar a influência de determinadas condições familiares do aluno e de algumas características do seu professor sobre o rendimento por ele obtido na disciplina de Matemática, assim como analisar as implicações que o aluno encara provocadas pela sua atitude em relação à Matemática. Pretende-se conhecer ainda a opinião dos alunos acerca da importância atribuída à Matemática na formação do indivíduo, para além de se procurar os factores que o aluno considera mais importantes para a aprendizagem da Matemática e os que mais contribuem para o insucesso desta disciplina. Mais especificamente, foram definidos como objectivos a alcançar:

- Identificar a associação entre a actividade profissional dos pais do aluno e o seu rendimento na disciplina de Matemática.

- Identificar a associação entre o grau de instrução dos pais do aluno e o seu rendimento na disciplina de Matemática.
- Identificar relações entre o enquadramento familiar do aluno, (nomeadamente se vive com os pais ou com outros familiares) e o seu rendimento na disciplina de Matemática.
- Identificar as relações entre o desempenho do aluno em Matemática em cada um dos anos do 3º ciclo e a capacidade pedagógica, científica e de relacionamento do seu professor, avaliadas pelo aluno.
- Inventariar os factores que, na visão dos alunos, mais contribuem para a aprendizagem da disciplina de Matemática.
- Inventariar os factores que, na visão dos alunos, mais contribuem para o insucesso da disciplina de Matemática.
- Analisar o percurso do rendimento do aluno na disciplina de Matemática ao longo dos três anos de escolaridade do terceiro ciclo.
- Identificar a associação entre o rendimento do aluno na disciplina de Matemática no final do terceiro ciclo e a escolha realizada para a área de estudos do ensino secundário.
- Caracterizar concepções e atitudes dos alunos em relação à Matemática.

Para conhecer o rendimento do aluno na disciplina de Matemática não foram usados quaisquer instrumentos de avaliação para além do nível obtido pelo aluno¹ no final de cada um dos três anos de escolaridade do 3º ciclo.

Com vista à concretização dos objectivos referidos, procedemos à elaboração de um questionário de opinião, instrumento este que constitui a base deste estudo. Conscientes da grande dimensão da população escolar que pretendemos conhecer - alunos a frequentar o 10º ano de escolaridade das Escolas Públicas do Distrito do Porto no ano lectivo de 1996/1997 - fomos levados à definição de uma amostra capaz de nos garantir uma dada precisão de resultados e um certo grau de confiança.

2.2 FASES DO ESTUDO

O presente estudo desenvolveu-se em diferentes fases: (i) definição da base de sondagem; (ii) concepção e redacção do questionário; (iii) pré-teste e ajuste do questionário; (iv) definição da amostra-piloto; (v) definição da dimensão da amostra final; (vi) distribuição dos questionários; (vii) recolha dos questionários; (viii) tratamento e análise de dados.

¹ Atendendo ao carácter de anonimato do questionário, foi considerado como real o nível que o aluno diz ter obtido, não tendo sido por isso confirmada essa informação. A veracidade destas respostas pode ser posta em dúvida, quer pela tendência que o aluno possa ter em não se desvalorizar, quer pelo facto da questão apelar à memória do aluno remetendo-o para três anos atrás.

2.2.1 DEFINIÇÃO DA BASE DE SONDAAGEM

A informação inicial necessária à constituição da base de sondagem pretendida foi facultada pela Direcção Regional de Educação do Norte (DREN), que disponibilizou "a Rede Escolar do Distrito do Porto para 96/97", ou seja, a listagem de todas as Escolas Públicas do Distrito do Porto que leccionaram o 10º ano de escolaridade no ano lectivo de 1996/1997, onde consta o número de turmas previsto nos cursos para prosseguimento de estudos (CSPOPE) e cursos para a vida activa (CSPOVA), em cada um dos quatro agrupamentos de estudo: Agrupamento 1 - Científica e Natural; Agrupamento 2 - Artes; Agrupamento 3 - Económico-Social; Agrupamento 4 - Humanidades.

Em seguida, contactamos todas as Escolas, no sentido de solicitarmos a informação relativa ao total de alunos em cada turma do 10º ano, por agrupamento de estudos.

Concluída a recolha desta informação, foi possível definir a base de sondagem que será o ponto de partida para a definição da amostra. Assim, no ano lectivo de 1996/1997 estavam inscritos nas 55 Escolas Públicas do Distrito do Porto que leccionaram o 10º ano de escolaridade um total de 18 947 alunos, distribuídos por 734 turmas.

Considerando a partição da população nos quatro agrupamentos atrás referidos, obtivemos a seguinte distribuição:

	Agrup 1	Agrup 2	Agrup 3	Agrup 4	Total
Nº alunos ²	9 075 (48%)	1 610 (8%)	3 710 (20%)	4 552 (24%)	18 947 (100%)
Nº turmas	343 (47%)	64 (9%)	150 (20%)	177 (24%)	734 (100%)

Tabela 2.1 Distribuição dos alunos e das turmas por agrupamento de estudos na população

2.2.2 CONCEPÇÃO E REDACÇÃO DO QUESTIONÁRIO

Depois de elaborarmos uma lista de questões que consideramos mais pertinentes no sentido dos objectivos definidos para este estudo, sentimos a necessidade de reduzir o número de questões, por nos parecer que seria excessivo e talvez muito demorado o tempo de resposta a essas questões.

A linguagem usada no texto das questões seleccionadas constituiu outra preocupação. Na redacção do questionário procuramos usar um vocabulário simples,

² Atendendo a que, durante o ano lectivo, o número de alunos de cada turma é variável, conforme vão ficando reprovados por faltas, anulam matrícula, desistem ou mudam de Escola, os números apresentados nesta Tabela correspondem ao total dos alunos fornecidos pelas Escolas no momento em que foi pedida a informação (Abril de 1997).

sem qualquer ambiguidade, por forma a que cada questão fosse perfeitamente clara e com o mesmo sentido para todos os alunos.

Ao optar pelas respostas de escolha múltipla na maior parte das questões formuladas, procuramos ainda asseguramo-nos de que nenhuma questão sugerisse qualquer resposta particular. Em três das questões, para além de propormos uma lista de possíveis respostas, demos a possibilidade de o aluno optar por uma resposta inteiramente sua e diferente das apresentadas.

Selecionadas as questões a introduzir, o passo seguinte foi o da escolha da sequência das questões que melhor resultaria. O encadeamento das questões foi assim cuidadosamente pensado, de modo a não haver mudanças bruscas de um assunto para outro.

O aspecto gráfico do questionário constituiu outra preocupação. O número de páginas foi outro aspecto a considerar, não só em termos orçamentais mas também de impacto para o aluno.

Assim, resultou um questionário de duas páginas, dividido em duas partes: a primeira, diz respeito a dados que identificam o aluno em termos de agrupamento, sexo, idade, agregado familiar e nível socio-económico dos pais do aluno; a segunda, relativo ao percurso do aluno na disciplina de Matemática no 3º ciclo do ensino básico.

2.2.3 DEFINIÇÃO DA AMOSTRA-PILOTO

Atendendo a que pretendemos fixar a precisão e o grau de confiança das estimativas, o procedimento a seguir seria o de definir uma amostra aleatória. Desta forma, na generalização dos resultados da amostra para a população, podemos controlar o erro amostral.

Ao optarmos por uma amostra aleatória, podemos ainda determinar qual a probabilidade de um aluno que frequenta o 10º ano de escolaridade no ano lectivo de 1996/1997 numa Escola Pública do Distrito do Porto ser seleccionado para a amostra, assegurando assim que ninguém é, *a priori*, excluído.

Por partirmos do pressuposto que o perfil do aluno, em termos de predisposição para a Matemática, seria diferente conforme o agrupamento de estudos que optara para prosseguir estudos no ensino secundário, pareceu-nos à partida que seria vantajoso construirmos uma amostra tendo por base a partição da população em quatro grupos, correspondentes aos agrupamentos de estudos atrás referidos.

Na definição da amostra é importante que esta capte a estrutura da população, isto é, que detenha as características que são fundamentais para o estudo. Acreditamos por isso que, optando pela amostra estratificada, estejamos a melhorar a representatividade da população. Por outro lado, a estratificação possibilitará a redução da dispersão da população, pelo facto de, neste tipo de amostragem, se criarem grupos que “variam muito entre si, no que diz respeito ao parâmetro em estudo, mas muito pouco dentro de si, ou seja, cada grupo é homogéneo e com pouca

variabilidade” (Vicente et al, 1996, p. 51). De facto, na estratificação acontece que a dispersão total da população se decompõe na soma da dispersão intra-estratos com a dispersão inter-estratos. Donde, em relação ao parâmetro em estudo, é desejável que, dentro de cada estrato os elementos sejam o mais semelhantes possível entre si, mas tão diferentes quanto possível dos elementos dos outros estratos.

Atendendo a que, por um lado, pretendíamos uma certa precisão dos resultados da amostra, e por outro, tencionávamos inferir esses mesmos resultados para a população, com determinado nível de confiança, precisávamos então determinar a dimensão da amostra a construir. Para tal, foi necessário estimar um parâmetro relativo à variável de referência em cada estrato.

Assim, tomaremos como variável de referência, X_{ij} , o aproveitamento na disciplina de Matemática³ no final do 9º ano de escolaridade, sendo $X_{ij} = 0$ se o j -ésimo aluno do estrato i obteve aproveitamento negativo (nível 1 ou 2) e $X_{ij} = 1$ se o j -ésimo aluno do estrato i obteve aproveitamento positivo (nível 3, 4 ou 5).

Pretendendo estimar a proporção p_i de alunos com aproveitamento positivo em Matemática no final do 3º ciclo, em cada um dos agrupamentos, facilmente se conclui que a estimação deste parâmetro toma por base uma população de Bernoulli.

Estimámos a proporção de alunos com aproveitamento positivo na disciplina de Matemática no final do 9º ano de escolaridade, a partir de quatro amostras-piloto, sendo cada uma constituída por 100 alunos pertencentes ao mesmo agrupamento. Considerando a dimensão da amostra, a variância σ^2 é estimada por $\hat{p}(1-\hat{p})$. Os resultados obtidos são os que constam no Tabela 2.2.

	Agrup 1	Agrup 2	Agrup 3	Agrup 4
\hat{p}_i	$\hat{p}_1=0.86$	$\hat{p}_2=0.73$	$\hat{p}_3=0.85$	$\hat{p}_4=0.71$
$\hat{\sigma}_i^2 = \hat{p}_i(1-\hat{p}_i)$	$\hat{\sigma}_1^2=0.1204$	$\hat{\sigma}_2^2=0.1971$	$\hat{\sigma}_3^2=0.1275$	$\hat{\sigma}_4^2=0.2059$
$\hat{\sigma}_i$	$\hat{\sigma}_1=0.0145$	$\hat{\sigma}_2=0.0388$	$\hat{\sigma}_3=0.0163$	$\hat{\sigma}_4=0.0424$

Tabela 2.2 Estimativa da proporção de alunos com aproveitamento positivo na disciplina de Matemática no final do 9º ano de escolaridade

2.2.4 DETERMINAÇÃO DA DIMENSÃO DA AMOSTRA

Tal como procedemos para a amostra-piloto, a amostra final foi definida considerando a partição da população nos quatro agrupamentos de estudos, cada um correspondendo a um estrato (Tabela 2.2). Cada aluno pertence a um e um só agrupamento, pelo que podemos afirmar que os estratos são mútua e exaustivamente exclusivos. A probabilidade de um aluno estar representado na amostra final será directamente proporcional à dimensão do agrupamento a que pertence, caso se opte por uma amostra proporcional, ou dependerá também da dispersão da variável em

³ Neste estudo designamos por “aproveitamento” a obtenção de nível superior a 2 no 3º Período.

estudo em cada um dos estratos, no caso de decidirmos pela repartição óptima de Neyman.

O efectivo global de sondagem n deve determinar-se de modo a satisfazer a condição

$$P(|\hat{p} - p| \leq \varepsilon) = 1 - \alpha.$$

A diferença entre o parâmetro e a estimativa não deve ser superior a ε (sendo este designado por precisão da estimativa) e $(1 - \alpha) \times 100\%$ representa o grau de confiança. Admitindo que ε é suficientemente pequeno e $N_i - n_i > 50$, prova-se que⁴

$$n = \frac{\sum_{i=1}^k (N_i^2 \sigma_i^2 / w_i)}{N^2 \left(\varepsilon / \chi_{\alpha/2} \right)^2 + \sum_{i=1}^k N_i \sigma_i^2} \quad (1)$$

onde N_i representa a dimensão do estrato i na população, n_i representa a dimensão do estrato i na amostra, N o total de indivíduos da população e w_i poderá tomar expressões diferentes conforme se opte por uma amostra estratificada proporcional ou pela amostra estratificada não proporcional (esta última conhecida por repartição óptima de Neyman). No primeiro caso temos

$$w_i = \frac{N_i}{N} \quad (2)$$

enquanto, no segundo, temos

$$w_i = \frac{N_i \sigma_i}{\sum_{i=1}^k N_i \sigma_i} \quad (3)$$

No que refere à precisão do estimador, a repartição óptima de Neyman é mais vantajosa em relação à estratificação proporcional, “no caso em que as dispersões σ_i são visivelmente diferentes ao longo dos estratos” (Gomes, 1998, p. 51).

Analisando a Tabela 2.2 no que respeita ao desvio padrão da variável em estudo em cada estrato, e atendendo a que $\hat{\sigma}_i$ toma o seu valor máximo quando $\hat{p}_i = 0.5$, ou seja, $0 \leq \hat{\sigma}_i \leq 0.5$, não podemos daí concluir se as diferenças entre os

⁴ Gomes, 1998, p. 53.

$\hat{\sigma}_i$ são ou não significativas. No entanto, sabemos que no caso em que as dispersões σ_i não diferem significativamente, “as duas abordagens fornecem resultados muito similares” (Gomes, 1998, p. 51). Assim sendo, optamos pela repartição óptima de Neyman, pelo que, para a determinação do efectivo global de sondagem, iremos recorrer à equação (1), onde w_i assume a expressão (3). Na amostragem estratificada não proporcional, quanto maior a dispersão do estrato maior será a proporção de indivíduos retirados, de modo a conseguir a representatividade da população em causa.

Fixando na expressão (3) o valor de ϵ em 0.02 e o grau de confiança em 95%, temos $n = 1306$

Conhecida a dimensão da amostra, procedemos à distribuição dos indivíduos pelos quatro estratos, de acordo com a repartição óptima de Neyman. Neste caso particular, em que não foram tomados em consideração os custos, quer o custo unitário de sondagem em cada estrato, quer os custos fixos de operação⁵, os efectivos de sondagem em cada estrato, n_i , são obtidos pela expressão

$$n_i = n \frac{N_i \sigma_i}{\sum_{i=1}^k N_i \sigma_i} \quad (4)$$

Desta expressão obtemos $n_1 \approx 567$, $n_2 \approx 129$, $n_3 \approx 238$ e $n_4 \approx 372$. A amostra será então constituída por 567 alunos do 1º agrupamento, 129 do 2º, 238 do 3º e 372 do 4º.

Dado que até à conclusão desta fase decorria já o mês de Maio de 1997, e com ele a proximidade do final do ano lectivo, havia que acelerar todo o processo de sondagem para que fosse possível a realização dos questionários antes dos alunos se encontrarem em período de férias. Por este motivo, procedemos à escolha aleatória de turmas em cada agrupamento (conglomerados os cachos) e em seguida inquirimos todos os alunos dessa turma, em vez de seleccionarmos os alunos aleatoriamente em cada agrupamento (o que exigia esforços incomportáveis, dada a grande dispersão das Escolas em estudo). Diminuímos assim significativamente o número de contactos a efectuar e consequentemente tornámos menos moroso todo o processo de sondagem.

Não podemos no entanto negligenciar a perda de precisão dos estimadores que resulta desta decisão. De facto, será previsível que haja uma certa homogeneidade entre os alunos de uma mesma turma no que respeita à variável em estudo, pelo que poderá surgir alguma redundância. Apesar disso, neste estudo concreto, um outro aspecto poderá contribuir para atenuar o efeito da homogeneidade dentro de cada turma: o facto do questionário procurar informação relativa ao 3º ciclo e no momento da sua aplicação os alunos estarem já no 10º ano. Na verdade, as turmas do 10º ano de escolaridade são, em geral, bastante heterogénias, por reunirem alunos que chegam de diferentes Escolas, que tiveram professores diferentes, utilizaram diferentes manuais escolares e até desconhecerem diferentes temas dos programas curriculares.

⁵ Todos os custos associados ao envio e recolha dos questionários foram suportados pela autora deste texto, pelo que não foram considerados na determinação dos efectivos de sondagem em cada estrato.

Determinámos o número de turmas, m_i , a inquirir no agrupamento i , dividindo o efectivo de sondagem de cada estrato, n_i , pela média de alunos por turma nesse estrato, N_{ic} . Deste modo, procurámos o valor de m_i de forma a cumprir a condição

$$m_i > \frac{n_i}{N_{ic}} \quad (5)$$

No Tabela 2.3 registámos o número de turmas e de alunos a inquirir, por forma a garantir um erro máximo amostral de 0,02 e um grau de confiança de 95% para a proporção de alunos com aproveitamento positivo na disciplina de Matemática no final do 3º ciclo.

	Agrup 1	Agrup 2	Agrup 3	Agrup 4	Total
Nº de alunos n_i	576 (43%)	131 (10%)	242 (18%)	378 (28%)	1327 (100%)
Nº de turmas m_i	29 (43%)	6 (9%)	13 (19%)	19 (28%)	67 (100%)

Tabela 2.3 Distribuição dos alunos e das turmas por agrupamento de estudos na amostra

Com o objectivo de seleccionarmos aleatoriamente as turmas a inquirir em cada agrupamento, numerámos as turmas e, recorrendo à folha de cálculo *EXCEL 97*, foram gerados números aleatórios.

Seleccionadas aleatoriamente as 67 turmas a inquirir, verificamos que os questionários iriam cobrir 34 Escolas de um total de 55 que constituem a base de sondagem.

2.2.5 ANÁLISE ESTATÍSTICA DOS DADOS

A análise estatística dos dados decorreu em três momentos:

- análise univariada (cálculo das medidas de tendências central e de dispersão para cada item quantitativo do questionário, construção de gráficos de frequências e *boxplots*);
- análise bivariada (construção de tabelas de frequências, medidas de associação, teste do qui-quadrado e análise da fiabilidade dos testes);
- análise multivariada (introdução ao método de AFCM, determinação do quadro de Burt, determinação dos maiores valores próprios, representação das modalidades em planos principais e interpretação dos eixos factoriais).

Nas secções seguintes abordaremos em pormenor cada uma destas etapas da análise estatística.

Neste trabalho abordaremos apenas os resultados da análise estatística univariada e multivariada.

3. ANÁLISE UNIVARIADA

As questões incluídas no questionário levaram à criação de variáveis e respectivas designações (mencionadas entre parêntesis e a negrito), que passaremos a apresentar:

- Agrupamento de estudos (**AGRUP**)
- Curso: Carácter Geral ou Tecnológico (**GERTEC**)
- Idade do aluno (**IDADE**)
- Sexo do aluno (**SEXO**)
- Habilitações literárias do pai do aluno (**HLP**)
- Habilitações literárias da mãe do aluno (**HLM**)
- Profissão do pai do aluno (**PP**)
- Profissão do mãe do aluno (**PM**)
- Pessoas com quem vive o aluno (**VIV**)
- O aluno teve o mesmo professor nos três anos do 3º ciclo (**MP3**)
- O aluno teve o mesmo professor em apenas dois anos (**MP2**)
- Anos com o mesmo professor (**MP789**)
- Capacidade de relacionamento do professor do 7º ano com os alunos (**CREL7**)
- Capacidade de relacionamento do professor do 8º ano com os alunos (**CREL8**)
- Capacidade de relacionamento do professor do 9º ano com os alunos (**CREL9**)
- Capacidade pedagógica do professor do 7º ano (**CPED7**)
- Capacidade pedagógica do professor do 8º ano (**CPED8**)
- Capacidade pedagógica do professor do 9º ano (**CPED9**)
- Capacidade científica do professor do 7º ano (**CC7**)
- Capacidade científica do professor do 8º ano (**CC8**)
- Capacidade científica do professor do 9º ano (**CC9**)
- Factores que mais contribuem para a aprendizagem da disciplina de Matemática:
 - Gosto pela Matemática (**API**)
 - Empenhamento do aluno (**AP2**)
 - Capacidade intelectual do aluno (**AP3**)
 - Interesse prático dos temas (**AP4**)
 - Adopção de um bom manual (**AP5**)
 - Utilização de diferentes materiais didácticos na aula (**AP6**)
 - Competência científica e pedagógica do professor (**AP7**)
 - Outros factores que contribuem para a aprendizagem da Matemática (**AP8**)
- Factores que mais contribuem para o insucesso da disciplina de Matemática:
 - Falta de conhecimentos básicos (**INS1**)
 - Desinteresse dos temas (**INS2**)
 - Falta de empenho dos alunos (**INS3**)
 - Disciplina muito abstracta (**INS4**)
 - Professores demasiado exigentes (**INS5**)
 - Manuais escolares pouco interessantes (**INS6**)
 - Monotonia das aulas (**INS7**)
 - Outros factores (**INS8**)
- Níveis obtidos pelo aluno no 3º período em cada um dos anos do 3º ciclo:
 - Nível do 7º ano (**N7**)
 - Nível do 8º ano (**N8**)
 - Nível do 9º ano (**N9**)
- No 10º ano o aluno tem: Matemática ou Métodos Quantitativos (**MATMQ**)
- Relação entre o aproveitamento em Matemática durante o 3º ciclo e a opção do agrupamento no 10º ano (**CONDP**)
- O aluno considera a Matemática importante para a sua formação (**FORM**)
- Razões que explicam a importância da disciplina de Matemática (**IMP**)
 - Desenvolve a capacidade de raciocínio e abstracção (**IMPI**)

- Facilita a interpretação da informação quantitativa (IMP2)
- É a base para a compreensão de outras ciências (IMP3)
- Tem aplicação em muitas actividades profissionais (IMP4)
- Está presente em diversas situações do nosso quotidiano (IMP5)
- Outros factores (IMP6)

Na **Tabela 3.1** apresentamos a classificação das variáveis de acordo com os valores por elas assumidos neste estudo.

Designação	Tipo de variável	Valores assumidos pela variável
AGRUP	Nominal	1 - Científica e Natural; 2 - Artes; 3 - Económico-Social; 4 - Humanidades
GERTEC	Dicotómica	1 - Geral; 2 - Tecnológico
IDADE	Discreta	Min = 14; Max = 22
SEXO	Dicotómica	0 - Feminino; 1 - Masculino
HLP	Discreta	Min = 1; Max = 30
HLM	Discreta	Min = 0; Max = 25
PP	Nominal	1 - Falecido; 2 - Reformado; 3 - Desempregado; outras ⁶
PM	Nominal	1 - Falecido; 2 - Reformado; 3 - Desempregada; 4 - Doméstica; outras ⁷
VIV	Nominal	0 - Não Respondeu; 1 - Pais; 2 - Só Mãe ou só Pai; 3 - Mãe e Padrasto ou Pai e Madrasta; 4 - Avós; 5 - outros
MP3	Dicotómica	0 - Não; 1 - Sim
MP2	Dicotómica	0 - Não; 1 - Sim
MP789	Nominal	78 - 7 ^o e 8 ^o anos; 79 - 7 ^o e 9 ^o anos; 89 - 8 ^o e 9 ^o anos
CREL7	Ordinal	1 - Pouca; 2 - Razoável; 3 - Muita
CREL8	Ordinal	1 - Pouca; 2 - Razoável; 3 - Muita
CREL9	Ordinal	1 - Pouca; 2 - Razoável; 3 - Muita
CPED7	Ordinal	1 - Pouca; 2 - Razoável; 3 - Muita
CPED8	Ordinal	1 - Pouca; 2 - Razoável; 3 - Muita
CPED9	Ordinal	1 - Pouca; 2 - Razoável; 3 - Muita
CC7	Ordinal	1 - Pouca; 2 - Razoável; 3 - Muita
CC8	Ordinal	1 - Pouca; 2 - Razoável; 3 - Muita
CC9	Ordinal	1 - Pouca; 2 - Razoável; 3 - Muita
AP1	Dicotómica	0 - Não; 1 - Sim
AP2	Dicotómica	0 - Não; 1 - Sim
AP3	Dicotómica	0 - Não; 1 - Sim
AP4	Dicotómica	0 - Não; 1 - Sim
AP5	Dicotómica	0 - Não; 1 - Sim
AP6	Dicotómica	0 - Não; 1 - Sim
AP7	Dicotómica	0 - Não; 1 - Sim
AP8	Dicotómica	0 - Não; 1 - Sim
INS1	Dicotómica	0 - Não; 1 - Sim
INS2	Dicotómica	0 - Não; 1 - Sim
INS3	Dicotómica	0 - Não; 1 - Sim
INS4	Dicotómica	0 - Não; 1 - Sim
INS5	Dicotómica	0 - Não; 1 - Sim
INS6	Dicotómica	0 - Não; 1 - Sim
INS7	Dicotómica	0 - Não; 1 - Sim

⁶ A codificação da profissão do Pai foi feita de acordo com a "Classificação Internacional de Profissões (versão portuguesa)"

⁷ A codificação da profissão da Mãe foi feita de acordo com a "Classificação Internacional de Profissões (versão portuguesa)"

INS8	Dicotómica	0 - Não; 1 - Sim
N7	Discreta	Min = 1; Max = 5
N8	Discreta	Min = 1; Max = 5
N9	Discreta	Min = 1; Max = 5
MATMQ	Dicotómica	0 - Métodos Quantitativos; 1 - Matemática
CONDP	Ordinal	1 - Pouco; 2 - Razoável; 3 - Muito
FORM	Dicotómica	0 - Não; 1 - Sim
IMP1	Dicotómica	0 - Não; 1 - Sim
IMP2	Dicotómica	0 - Não; 1 - Sim
IMP3	Dicotómica	0 - Não; 1 - Sim
IMP4	Dicotómica	0 - Não; 1 - Sim
IMP5	Dicotómica	0 - Não; 1 - Sim
IMP6	Dicotómica	0 - Não; 1 - Sim

Tabela 3.1 Classificação das variáveis segundo a escala em que são expressas

Começaremos por estudar cada uma das variáveis intervenientes neste estudo, ou seja, procuraremos descrever a amostra pondo em evidência as suas principais características.

Assim, na Tabela 3.2 apresentamos os valores da média, desvio padrão e quartis das variáveis quantitativas em estudo.

Em relação à idade dos estudantes a frequentar o 10º ano de escolaridade, atendendo a que entram para o 1º ano de escolaridade com 6 anos, seria de esperar de tivessem nesta altura 16 anos (admitindo, naturalmente, não ter ocorrido nenhuma retenção até então).

	N	Média	Desvio Padrão	1º Quartil	2º Quartil	3º Quartil
IDADE	1315	16,06	0,99	15	16	17
HLP	1185	7,67	4,74	4	5	11,5
HLM	1206	7,30	4,57	4	4	10
N7	1287	3,50	0,92	3	3	4
N8	1292	3,38	0,95	3	3	4
N9	1296	3,32	0,95	3	3	4

Tabela 3.2 Medidas descritivas das variáveis quantitativas em estudo

Podemos verificar no Gráfico 3.1 que a maior parte dos alunos tem de facto 16 anos (44%), existindo, no entanto, um número significativo de alunos ainda com 15 anos e outros já com 17. Obviamente que esta situação poderá acontecer por diversas razões, uma das quais será a data de aniversário⁸ e outra, no caso dos alunos com 17 anos, eventualmente a ocorrência de uma retenção no percurso escolar do aluno. Observamos ainda que cerca de 9% dos alunos inquiridos têm mais de 17 anos, situação que corresponderá por certo a retenções em anos anteriores.

⁸ Não foi objecto de estudo desta investigação.

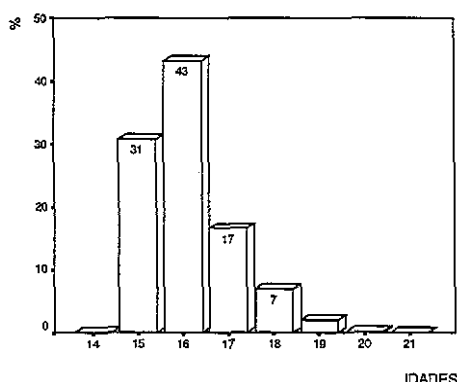


Gráfico 3.1 Distribuição dos alunos por idades

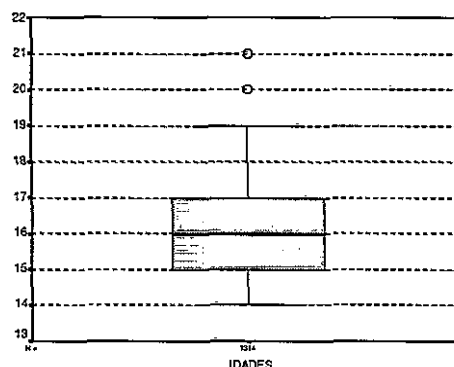


Gráfico 3.2 Boxplot relativo às idades dos alunos

Pela análise do Gráfico 3.2 podemos constatar que pelo menos 50% dos alunos inquiridos têm idades entre 15 e 17 anos, inclusive, havendo alguns valores considerados *outliers*. Neste caso, 20 e 21 anos são valores invulgares na distribuição das idades. Segundo Bento Murteira (1993), estes valores são designados por *outliers* moderados por pertencerem ao intervalo

$$[Q_{75} + 1.5 (Q_{75} - Q_{25}), Q_{75} + 3 (Q_{75} - Q_{25})]$$

ou seja,

$$[17 + 1.5 (17-15), 17 + 3 (17-15)] = [20, 23]$$

Sendo a média um parâmetro muito sensível a estes valores extremos, tomaremos a mediana (16 anos) como medida que melhor caracteriza o conjunto das idades. Porém, atendendo à grande dimensão da amostra (1327 indivíduos), seria de esperar que os valores da média e da mediana fossem muito próximos, como de facto aconteceu (16,07 ≈ 16).

Observando a Tabela 3.2 é-nos permitido afirmar que, em média, os pais possuem o 7º ano de escolaridade. Mas esta informação é, por si só, enganosa, uma vez que, quer a variável HLP quer a HLM, têm um desvio padrão considerável (5,56 e 5,49, respectivamente). Por outro lado, os Gráficos 3.3 e 3.4 evidenciam a elevada percentagem de pais (cerca de 48%) que concluíram apenas o 1º ciclo (apesar de, no nosso país, desde de 1991, a escolaridade obrigatória ser de nove anos). Podemos constatar também que abandonaram a Escola depois de concluírem o 2º ciclo (sendo este o anterior limite de obrigatoriedade escolar) aproximadamente 9% dos pais e 10% das mães dos alunos inquiridos. Não prosseguiram estudos para além do 12º ano de escolaridade cerca de 9% dos pais e 8% das mães. Encontramos apenas 11% das mães e 13% dos pais que terminaram um curso superior.

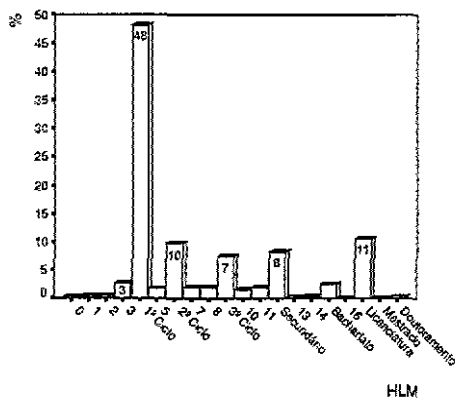


Gráfico 3.3 Habilitações literárias do pai do aluno

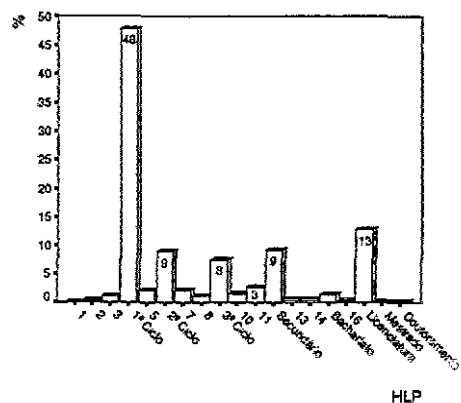


Gráfico 3.4 Habilitações literárias da mãe do aluno

Pela análise da *boxplot* relativa a estas duas variáveis (Gráfico 3.5), facilmente se conclui que as mães dos alunos inquiridos têm, de um modo geral, menos habilitações literárias relativamente aos pais dos mesmos alunos.

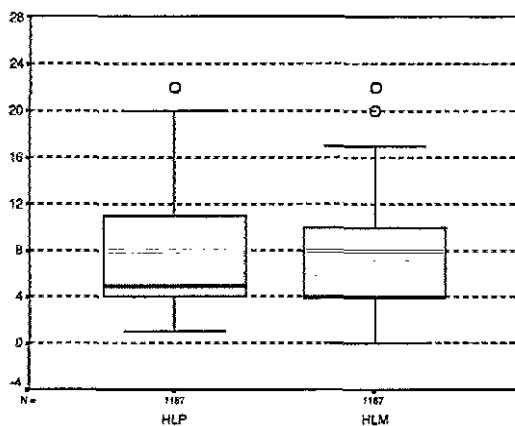


Gráfico 3.5 *Boxplot* relativo às habilitações literárias do pai e da mãe do aluno

Por outro lado, o Gráfico 3.5 dá-nos ainda informação acerca dos anos de escolaridade que na amostra surgem como *outliers*. Assim, verificamos que a formação depois da Licenciatura (HLP e HLM superiores a 17) ocorre apenas excepcionalmente no caso das mães dos alunos. Os pais Mestres não são considerados *outliers*, enquanto os Doutorados são de facto casos muito raros no universo em estudo, recebendo também a designação de *outliers*.

No que respeita ao aproveitamento na disciplina de Matemática, o balanço, ao contrário do que tantas vezes se comenta na comunicação social, pode considerar-se francamente bom. Note-se que pelo menos 44% dos alunos atingiram pelo menos o nível 3, em qualquer dos três anos, sendo importante registarmos que pelo menos 38% dos alunos obtiveram nível 4 ou 5 (Gráficos 3.6, 3.7 e 3.8). A percentagem de níveis negativos é cerca de 11 no 7º ano, parecendo, no entanto, haver uma tendência para um crescimento do insucesso em Matemática, à medida que o aluno avança no 3º ciclo.

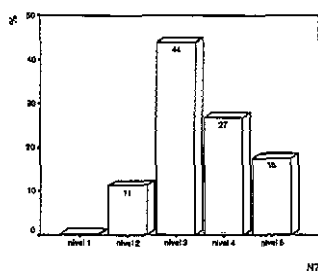


Gráfico 3.6 Nível obtido no 7º ano na disciplina de Matemática

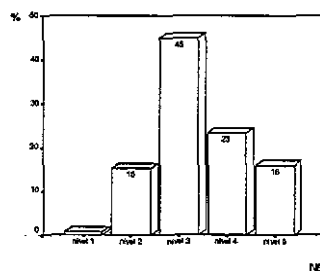


Gráfico 3.7 Nível obtido no 8º ano na disciplina de Matemática

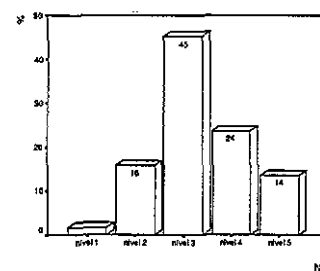


Gráfico 3.8 Nível obtido no 9º ano na disciplina de Matemática

A *boxplot* referente ao aproveitamento em Matemática no percurso do 3º ciclo (Gráfico 3.9) evidencia, por um lado, o nível 1 como valor excepcional em qualquer dos três anos e, por outro, a semelhança de resultados nos diferentes anos (como aliás era de prever pelos resultados da Tabela 2.1).

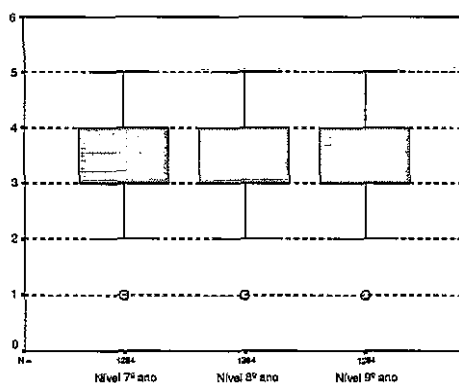


Gráfico 3.9 *Boxplot* relativo ao aproveitamento em Matemática durante o 3º ciclo

Tal como prevê o artigo 10º da Lei de Bases do Sistema Educativo, o ensino secundário organiza-se em dois cursos, um orientado para a vida activa (curso tecnológico) e outro para o prosseguimento de estudos (curso de carácter geral). Assim, verificamos que a preferência dos alunos inquiridos vai para os cursos de carácter geral, ou seja, será intenção da maioria dos alunos (77%) prosseguir os seus estudos para ingresso no ensino superior (Gráfico 3.11).

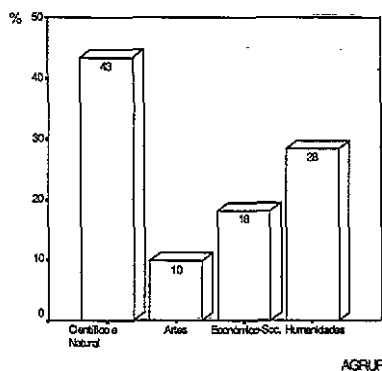


Gráfico 3.10 Distribuição dos alunos por agrupamentos

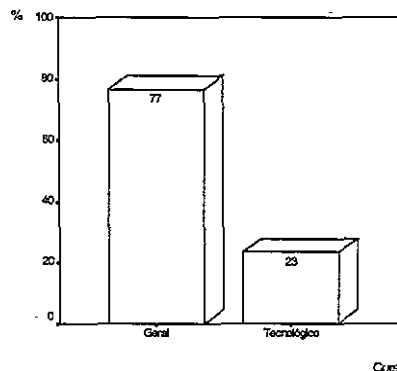


Gráfico 3.11 Distribuição dos alunos por cursos

A distribuição dos alunos inquiridos pelos quatro agrupamentos de estudos difere um pouco da população em estudo, pelo facto de na amostra termos optado pela

repartição óptima de Neyman, que retira tantos mais indivíduos quanto maior for a dispersão do estrato. Comparando o Gráfico 3.10 com a Tabela 2.1, verificamos que a amostra tem menos 5% de alunos no agrupamento 1, enquanto no agrupamento 4 tem mais 4%; por outro lado, tem menos 2% no agrupamento 3 e mais 1% no agrupamento 2.

De notar que, apesar do agrupamento 1 - Científica e Natural - exigir a frequência da disciplina de Matemática até ao 12º ano como disciplina específica, a preferência dos alunos é maioritariamente por este agrupamento (48%). Segue-se o agrupamento de Humanidades (24%), que apenas tem a disciplina de Métodos Quantitativos no 10º ou 11º anos. O agrupamento 3 - Económico-Social - que tem também a disciplina de Matemática como obrigatória nos 3 anos do curso secundário, é frequentado por 20% dos alunos do ensino público do distrito do Porto. O curso de Artes é o menos procurado (8%), sendo este agrupamento o único em que o aluno pode optar pela disciplina de Métodos Quantitativos de frequência apenas num ano ou pela disciplina de Matemática a frequentar em 3 anos lectivos do ensino secundário.

Pela análise do Gráfico 3.12 verificamos que no Distrito do Porto, e ao contrário do que acontecia à alguns anos atrás⁹, há um maior número de raparigas a estudar em relação ao número de rapazes.

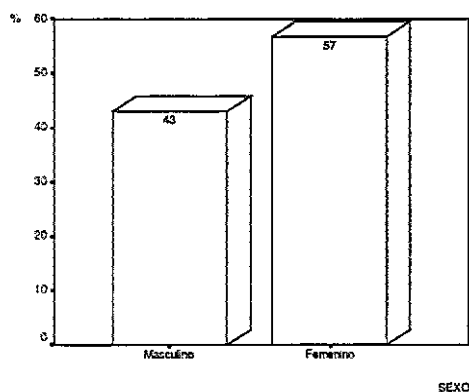


Gráfico 3.12 Distribuição dos alunos por sexo

Sendo um dos objectivos do questionário a caracterização por parte dos alunos dos seus professores de Matemática do 3º ciclo, em termos pedagógicos, científicos e de capacidade de relacionamento com os alunos, era importante sabermos se houve ou não continuidade do professor durante o ensino-aprendizagem nesse ciclo. Assim, verificamos que apenas 15 % dos alunos dizem ter tido o mesmo professor nos 3 anos do 3º ciclo, enquanto que, dos restantes, cerca de 22 % afirmam ter tido o mesmo professor em apenas 2 desses anos (Gráficos 3.13, 3.14 e 3.15).

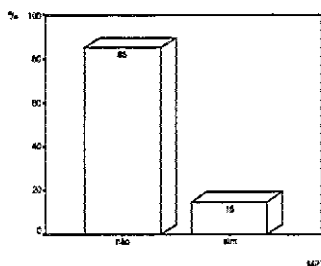


Gráfico 3.13 O aluno tem o mesmo professor no 7º, 8º e 9º ano?

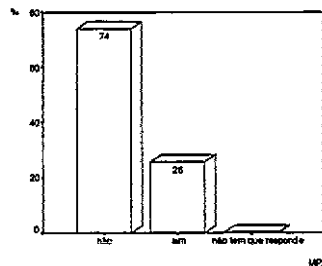


Gráfico 3.14 O aluno tem o mesmo professor em dois anos?

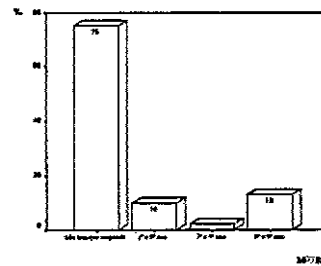


Gráfico 3.15 Anos em que o aluno teve o mesmo professor

⁹ Como facilmente se pode comprovar pelas Estatísticas da Educação do INE.

A opinião pública sobre o professor da disciplina de Matemática nem sempre é a mais favorável. No entanto, podemos verificar que os alunos inquiridos consideram ser, na generalidade dos casos, razoável ou boa a capacidade de relacionamento do professor com os seus alunos (Gráficos 3.16, 3.17 e 3.18). Registe-se entretanto que oscila entre 14% a 19% os inquiridos que consideram que tiveram professores de Matemática com pouca capacidade de relacionamento com os seus alunos, em cada um dos anos que constitui o 3º ciclo.

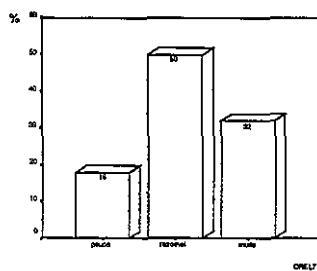


Gráfico 3.16 Capacidade de relacionamento do professor do 7º ano

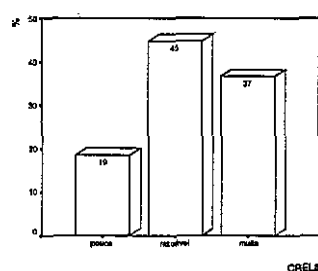


Gráfico 3.17 Capacidade de relacionamento do professor do 8º ano

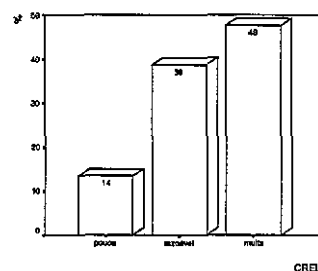


Gráfico 3.18 Capacidade de relacionamento do professor do 9º ano

No que respeita à capacidade pedagógica do professor de Matemática, mais de 11 % dos alunos afirmam ser deficiente, embora mais de 33 % dos alunos digam ter tido professores desta disciplina com muita capacidade pedagógica (Gráficos 3.19, 3.20 e 3.21).

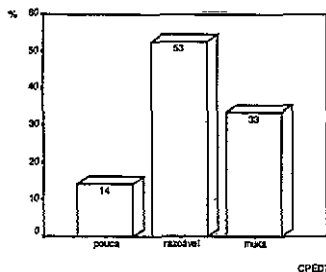


Gráfico 3.19 Capacidade pedagógica do professor do 7º ano

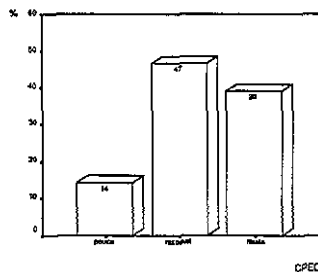


Gráfico 3.20 Capacidade pedagógica do professor do 8º ano

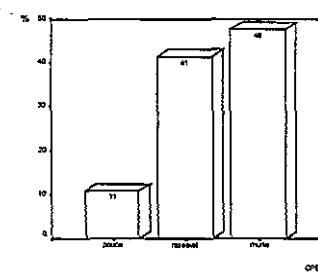


Gráfico 3.21 Capacidade pedagógica do professor do 9º ano

Situação semelhante acontece em relação à capacidade científica dos professores de Matemática dos alunos inquiridos. Estes consideram maioritariamente que os seus professores são cientificamente razoáveis ou muito bons, registando-se, no entanto, nos diferentes anos, uma margem de 9 a 13 % dos alunos a considerarem os respectivos professores com pouca capacidade científica (Gráficos 3.22, 3.23 e 3.24).

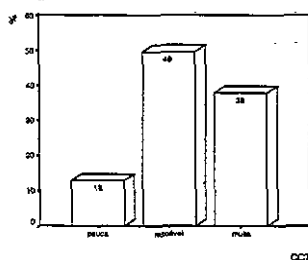


Gráfico 3.22 Capacidade científica do professor de 7º ano

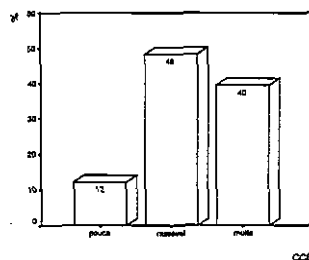


Gráfico 3.23 Capacidade científica do professor de 8º ano

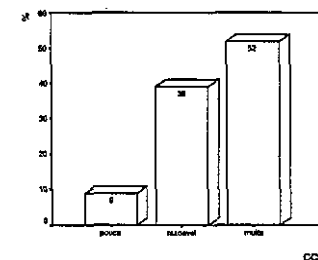


Gráfico 3.24 Capacidade científica do professor de 9º ano

Sendo a disciplina de Matemática obrigatória para uns alunos e de opção (entre Matemática e Métodos Quantitativos) para outros, conforme o agrupamento a que pertencem, foi pedido aos alunos que dissessem qual a disciplina que tinham no 10º ano. O Gráfico 3.25 permite concluir, numa leitura rápida, que apenas 38% têm a disciplina de Métodos Quantitativos¹⁰. Se pensarmos que 28% dos inquiridos pertencem ao 4º agrupamento e como tal não têm Matemática no ensino secundário, significa então que os restantes 10% são os alunos do 2º agrupamento que optaram todos por Métodos Quantitativos, apesar das limitações que este facto lhes acarreta por verem bastante reduzido o leque de escolhas em termos cursos superiores.

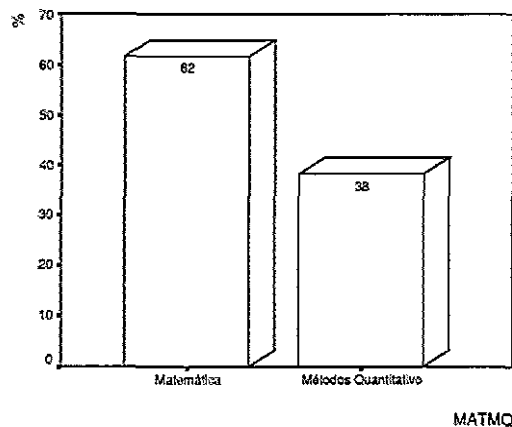


Gráfico 3.25 Disciplina no 10º ano

Com o intuito de sabermos quais os factores que os alunos consideram mais contribuir para a aprendizagem da disciplina de Matemática, enumeramos uma lista de 7 possíveis factores, deixando em aberto a possibilidade do aluno apresentar um factor diferente, pedimos que escolhesse quatro deles. No Gráfico 3.26 podemos observar quais os factores indicados mais frequentemente de entre as 8 possibilidades apresentadas, destacando-se o “empenhamento do aluno” (AP2), o “gosto pela Matemática” (AP1), logo seguido da “competência pedagógica e científica do professor” (AP7). O factor que parece ter merecido menos importância foi a “utilização de diferentes materiais didácticos na aula” (AP6), com apenas 19% dos alunos a seleccionar esta opção.

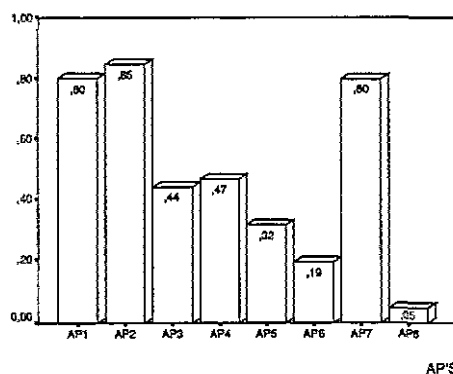


Gráfico 3.26 Factores que mais contribuem para a aprendizagem da disciplina de Matemática

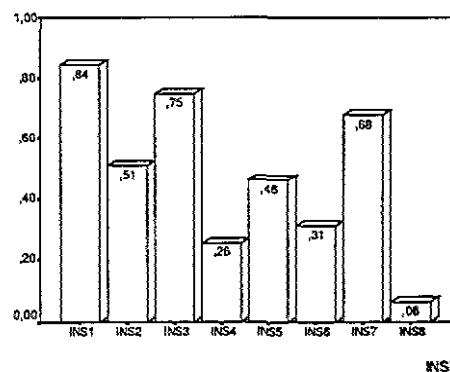


Gráfico 3.27 Factores que mais contribuem para o insucesso da disciplina de Matemática

¹⁰ Notemos que esta disciplina, nomeadamente para os alunos do agrupamento 4, é obrigatória.

Se pensamos ser de extrema importância conhecer os factores que mais contribuem, segundo a óptica dos alunos, para a aprendizagem da disciplina de Matemática, importante será também conhecer as razões que estes atribuem para o tão comentado insucesso da Matemática. O Gráfico 3.27 dá informação sobre as preferências dos alunos para justificarem tal situação. Foi apresentada, de modo análogo, uma lista com 8 factores (sendo a última opção aberta) e foram pedidos 4 destes factores. Destacamos o motivo mais apontado, “a falta de conhecimentos básicos” (INS1), seguido da “falta de empenho dos alunos” (INS3) e da “monotonia das aulas” (INS7). Outras razões também evocadas (embora com menor ênfase) para explicarem o insucesso da Matemática, foram o “desinteresse dos temas” (INS2) e “professores demasiado exigentes” (INS5).

Verificamos também que cerca de 30% dos alunos inquiridos consideram que a Matemática não tem qualquer interesse para a sua formação (Gráfico 3.29). Por outro lado, 26% dos alunos assumem que o seu aproveitamento na disciplina de Matemática durante o 3º ciclo condicionou a sua opção do agrupamento de estudos para o 10º ano, enquanto sensivelmente outros tantos afirmam não terem sido nada influenciados por aquele aspecto (Gráfico 3.28).

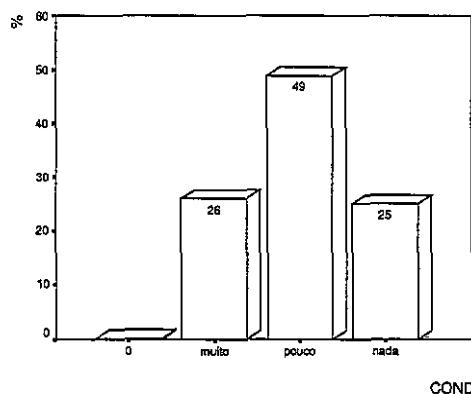


Gráfico 3.28 Os resultados obtidos no 3º ciclo condicionaram a opção do agrupamento de estudos no 10º ano?

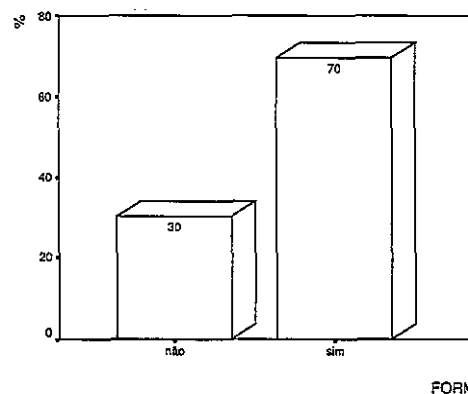


Gráfico 3.29 Considera a Matemática importante para a sua formação?

Por certo, alguma preocupação ressalta pelo facto de 75% dos alunos inquiridos reconhecerem a interferência do seu rendimento na disciplina de Matemática nas suas decisões de escolha de opção relativamente ao agrupamento de estudo do ensino secundário e consequentemente para o ensino superior.

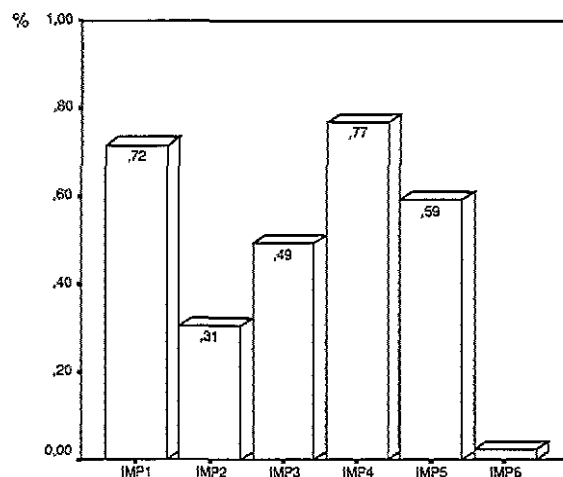


Gráfico 3.30 Razões que explicam a importância atribuída pelo aluno à disciplina de Matemática

A última pergunta do questionário destina-se apenas aos alunos, que na questão anterior, dizem reconhecer a importância da Matemática para a sua formação. Perante um conjunto de 6 opções, é pedido que o aluno seleccione as 3 que melhor explicam a importância que ele próprio atribui à disciplina de Matemática. Assim, como podemos observar no Gráfico 3.30, a quarta opção “tem aplicação em muitas actividades profissionais” foi a mais evocada, logo seguida da primeira opção “desenvolve a capacidade de raciocínio e abstracção” (com 77% e 72%, respectivamente).

4. ANÁLISE MULTIVARIADA

4.1 INTRODUÇÃO

Nas secções anteriores estudámos as variáveis individualmente e duas a duas. Contudo, tratando-se da análise de um conjunto considerável de variáveis, só será possível a clarificação da estrutura subjacente com a análise simultânea de todas as variáveis. Os métodos de estatística descritiva multidimensional permitem que se explore o desempenho conjunto das variáveis e se determine a importância de cada uma em presença das restantes. O objectivo essencial destes métodos, desenvolvidos particularmente nas três últimas décadas em França, é o de “fornecer resumos da informação contida em grandes quadros de dados envolvendo essencialmente variáveis quantitativas e qualitativas” (Gomes, 1995, p.13).

A falta de cálculo automático apropriado até à década de sessenta levou a que a utilização das diferentes técnicas da estatística multivariada fosse muito limitada. Os efeitos do progresso tecnológico fez-se sobretudo sentir na rapidez com que se analisa uma base de dados de grande dimensão e relativa complexidade. Assim, os cientistas ficaram mais libertos para se lançarem em outro tipo de actividades, nomeadamente no desenvolvimento de novos modelos de análise teórica.

A leitura dos gráficos onde se representam os indivíduos permite que se estabeleçam semelhanças ou diferenças entre eles, enquanto que a representação das variáveis evidencia eventuais associações entre estas.

Na estatística multivariada, tal como na univariada e na bivariada, existem métodos exploratórios, cujo principal objectivo é descrever as características de um conjunto de dados, e métodos confirmatórios, cujo pressuposto da normalidade multivariada da população de onde foram retiradas as observações é essencial, de forma a permitir a inferência dos resultados da amostra para a população. Embora esta divisão seja conveniente em termos pedagógicos, na prática os dois métodos são muitas vezes aplicados em simultâneo.

Neste estudo, de entre os vários métodos da análise multivariada, iremos utilizar a análise factorial das correspondências múltiplas (AFCM), método privilegiado para tratamento de dados resultantes de questionários, onde predominam variáveis qualitativas.

Procurámos usar para esta análise um *software* estatístico actual e com as potencialidades pretendidas. Daí a nossa escolha ter recaído sobre o *SPAD* versão 3.0, lançado no mercado em 1997.

Procederemos inicialmente a uma breve referência a alguns aspectos teóricos necessários a uma melhor compreensão deste tipo de abordagem.

Consideremos o quadro inicial, designado por quadro disjuntivo completo, constituído por n linhas (indivíduos) e R colunas (total de modalidades de resposta do questionário), descrevendo as T respostas dos n indivíduos. Seja R_j o número de modalidades de resposta da questão j (com $j = 1, 2, \dots, T$). Temos assim,

$$R = \sum_{j=1}^T R_j$$

O quadro disjuntivo completo Z é um quadro binário, onde é pressuposto que o indivíduo i escolha apenas uma modalidade em cada uma das T questões. Temos então,

$$Z = [Q_1 | Q_2 | \dots | Q_T], \text{ ou seja,}$$

Considerando o sub-quadro Q_j , verificamos que ele é constituído por n linhas e R_j colunas, e que a i -ésima linha contém $(R_j - 1)$ zeros, correspondendo o valor 1 à modalidade da questão j seleccionada pelo indivíduo i .

Por vezes os dados são apresentados num quadro “condensado”, reduzindo o número colunas de R para T . Assim, cada valor (i, j) deste quadro representa o número da modalidade da questão j escolhida pelo indivíduo i .

Pode haver interesse em trabalhar com um outro quadro associado a Z , conhecido como tabela de contingência generalizada ou quadro de Burt, que resulta do cruzamento de todas as T respostas entre si. O quadro de Burt, B , obtém-se fazendo $B = {}^t Z Z$, tem T^2 blocos sendo a soma dos elementos em cada bloco é igual a n .

Quer se considere este quadro B de ordem (R, R) ou o quadro Z de ordem (n, R), prova-se que a análise factorial das correspondências resultante é equivalente¹¹.

Realizar uma AFC de um destes quadros consiste em determinar os factores principais resultantes da diagonalização da matriz

$$\frac{1}{T}(\text{diag}B)^{-1}B \quad \text{se partirmos do quadro Z (quadro disjuntivo completo),}$$

ou

$$\frac{1}{T^2}(\text{diag}B)^{-1}{}^tB(\text{diag}B)^{-1}B \quad \text{se partirmos do quadro B (quadro de}$$

Burt)¹².

Desta forma, pretende-se caracterizar os indivíduos através da “informação principal” contida na matriz inicial. A análise factorial das correspondências consiste na resolução de um problema de optimização, na medida em que o seu principal objectivo é “reconstituir o máximo da variabilidade existente nos dados iniciais através de um número reduzido de novas variáveis” (Gomes, 1993, p. 21). Estas variáveis, designadas por componentes principais, resultam da diagonalização de uma das matrizes acima referidas, conforme $n < R$ ou $n > R$, respectivamente. O SPAD procura a matriz de menor dimensão, ou seja, a matriz mais rentável em termos de diagonalização e, a partir dela, determina os q maiores valores próprios, λ_k , e os vectores próprios associados, u_k . Dito de outro modo, a AFC pretende projectar os n indivíduos pertencentes a um espaço de dimensão R num sub-espaço de dimensão q (com $q < R$), de forma que a deformação global da projecção dos indivíduos seja mínima.

Um indicador global da qualidade da AFCM é a percentagem de inércia explicada pelos q primeiros eixos. A inércia total da nuvem de pontos, I_g , é uma medida da dispersão global dos pontos, donde, por definição, é a média dos quadrados das distâncias dos vectores $x_{\sim j}$ do quadro de perfil de coluna associados às

R modalidades presentes e o vector g centro de gravidade dos vectores $x_{\sim j}$. Assim,

$$I_g = \sum_{j=1}^R f_j \left\| x_{\sim j} - g \right\|_{D^{-1}}^2$$

sendo a métrica $D^{-1} = \begin{pmatrix} n & 0 & \dots & 0 \\ 0 & n & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & n \end{pmatrix}$ a matriz diagonal cujos elementos

diagonais são os inversos dos pesos dos indivíduos.

Considerando a análise a partir do quadro Z, de dimensão (n, R), verificamos que a soma de todos os seus elementos é nT , pois cada um dos n indivíduos deve

¹¹ A demonstração deste teorema pode ser encontrada em Gomes, 1993, p. 141.

¹² A dedução destas expressões pode encontrar-se em Gomes, 1993

escolher apenas uma modalidade de cada uma das T questões. Assim, designando por n_j a soma dos elementos da coluna j desta matriz, temos $f_j = \frac{n_j}{nT}$. Pode então escrever-se

$$I_g = \sum_{j=1}^R \frac{n_j}{nT} {}^t \begin{pmatrix} x_j - g \\ \sim_j \end{pmatrix} D^{-1} \begin{pmatrix} x_j - g \\ \sim_j \end{pmatrix}$$

Mas, o vector $x_{\sim_j} = \begin{pmatrix} X_{1j}/n_j \\ X_{2j}/n_j \\ \vdots \\ X_{nj}/n_j \end{pmatrix}$, onde

$$X_{ij} = \begin{cases} 1 & \text{se o indivíduo } i \text{ escolheu a modalidade } j \\ 0 & \text{senão} \end{cases} \quad \text{e} \quad g_{\sim} = \begin{pmatrix} 1/n \\ 1/n \\ \vdots \\ 1/n \end{pmatrix}$$

Donde,

$$I_g = \sum_{j=1}^R \frac{n_j}{nT} n \begin{pmatrix} \frac{X_{1j}}{n_j} - \frac{1}{n} & \frac{X_{2j}}{n_j} - \frac{1}{n} & \dots & \frac{X_{nj}}{n_j} - \frac{1}{n} \end{pmatrix} \begin{pmatrix} \frac{X_{1j}}{n_j} - \frac{1}{n} \\ \frac{X_{2j}}{n_j} - \frac{1}{n} \\ \vdots \\ \frac{X_{nj}}{n_j} - \frac{1}{n} \end{pmatrix}$$

$$I_g = \frac{1}{T} \sum_{j=1}^R n_j \sum_{i=1}^n \left(\frac{X_{ij}}{n_j} - \frac{1}{n} \right)^2 \quad (6)$$

Desenvolvendo o quadrado e reconhecendo que $\sum_{i=1}^n X_{ij} = \sum_{i=1}^n X_{ij}^2 = n_j$,

temos que

$$I_g = \frac{1}{T} \sum_{j=1}^R \left(1 - \frac{n_j}{n} \right) = \frac{1}{T} \left(R - \frac{\sum_{j=1}^R n_j}{n} \right) = \frac{1}{T} (R - T) = \frac{R}{T} - 1$$

Podemos assim afirmar que a inércia total é igual ao valor médio de modalidades em estudo diminuído de uma unidade.

Por outro lado, a inércia total corresponde numericamente à soma dos valores próprios não triviais¹³, ou seja, ao traço da matriz a diagonalizar diminuído de uma unidade. Pode facilmente verificar-se que, realizando a AFCM partindo do quadro disjuntivo completo, Z, vem

$$I_g = \text{Tr} \left[\frac{1}{T} (\text{diag } B)^{-1} B \right] - 1 = \text{Tr} \left(\frac{1}{T} I_R \right) - 1 = \frac{1}{T} R - 1$$

(onde I_R representa a matriz identidade de ordem R)

resultado este que obtivemos atrás, partindo da definição de inércia total. Podemos então escrever,

$$I_g = \sum_{k=1}^{\gamma} \lambda_k = \frac{R}{T} - 1 \quad (\text{com } \gamma \text{ igual ao número de valores próprios não trivialmente iguais a 0 ou a 1})$$

Verifica-se ainda que, numa AFCM, o número de valores próprios, não trivialmente iguais a 0 ou a 1, é igual à diferença entre o total de modalidades e o total de questões, isto é, $R - T$. Assim sendo, facilmente se conclui que a média dos valores próprios não trivialmente iguais a 0 ou 1 é

$$\frac{\sum_{k=1}^{\gamma} \lambda_k}{R - T} = \frac{\frac{R}{T} - 1}{R - T} = \frac{1}{T}$$

Este resultado é vulgarmente usado como critério para a selecção dos q maiores valores próprios numa AFCM, rejeitando-se os que verificam a condição $\lambda_k < \frac{1}{T}$.

Pretendendo-se calcular a distância entre a j -ésima coluna do quadro disjuntivo completo e o centro de gravidade da nuvem, temos

$$d^2 \left(\begin{matrix} x_j, g \\ \sim \quad \sim \end{matrix} \right) = n \sum_{i=1}^n \left(\frac{X_{ij}}{n_j} - \frac{1}{n} \right)^2$$

onde X_{ij} é o i -ésimo termo da j -ésima coluna do quadro disjuntivo completo e n_j é a soma dos n elementos da coluna j (número de indivíduos que escolheu a modalidade j).

Esta expressão que resulta de se tomar a j -ésima parcela da expressão (6), pode escrever-se,

$$d^2 \left(\begin{matrix} x_j, g \\ \sim \quad \sim \end{matrix} \right) = \frac{n}{n_j} - 1$$

¹³ Gomes, 1993, p. 117

Assim, a inércia associada a essa modalidade j é dada pela expressão

$$\frac{n_j}{nT} d^2 \left(\begin{matrix} x_j, g \\ \sim \quad \sim \end{matrix} \right) = \frac{1}{T} \left(1 - \frac{n_j}{n} \right)$$

e a inércia associada a uma dada questão com R_k modalidades é

$$\sum_{j=1}^{R_k} \frac{1}{T} \left(1 - \frac{n_j}{n} \right) = \frac{R_k - 1}{T}$$

A contribuição dessa questão para a inércia total é

$$\frac{\frac{R_k - 1}{T}}{\frac{1}{T} \sum_{K=1}^T R_k - 1} = \frac{R_k - 1}{\sum_{K=1}^T (R_k - 1)}$$

A percentagem de inércia explicada pelo eixo k é

$$\frac{\lambda_k}{\frac{1}{T} \sum_{K=1}^T R_k - 1}$$

Sabendo que numa AFC qualquer valor próprio é um valor entre 0 e 1 inclusive, temos

$$\frac{\lambda_k}{\frac{1}{T} \sum_{K=1}^T R_k - 1} < \frac{1}{\frac{R}{T} - 1}$$

isto é, um limite superior para a percentagem de inércia explicada pelo eixo k .

A contribuição absoluta da modalidade j para a formação do k -ésimo eixo principal de inércia é

$$CTA_j^k = \frac{f_j (y_j^k)^2}{\lambda_k}$$

com $j \in Q_j$, sendo $f_j = \frac{n_j}{nT}$ e Y_j^k a coordenada da modalidade j projectada no k -ésimo eixo principal.

Consequentemente, a contribuição absoluta da questão Q_i para a formação do k -ésimo eixo principal de inércia é

$$CTA_{Q_i}^k = \frac{\sum_{j=1}^{R_i} f_j (Y_j^k)^2}{\lambda_k}$$

A contribuição relativa do plano $\begin{bmatrix} u_{\sim k} & u_{\sim k'} \end{bmatrix}$ à modalidade j , isto é, a qualidade de representação da projecção do ponto de IR^n associado a essa modalidade é

$$\frac{(Y_{J_j}^k)^2 + (Y_{J_j}^{k'})^2}{\sum_{i=1}^n \frac{1}{f_i} \left(\frac{f_{ij}}{f_j} - f_i \right)^2}$$

4.2 APLICAÇÃO DA AFCM AO QUESTIONÁRIO EM ESTUDO

A importação dos dados para o *SPAD* exige que estes tenham sido previamente gravados com extensão .txt . Depois de termos importado a base de dados de *EXCEL* para *SPAD*, sentimos a necessidade de fazer nova recodificação das variáveis, pois nesta altura verificámos existirem algumas incompatibilidades de um *software* para outro, nomeadamente no que se refere à designação para os valores omissos, que até então convencionámos ser 99 e em *SPAD* é obrigatoriamente 0, podendo assumir um valor diferente apenas no caso das variáveis contínuas. Por outro lado, o valor 0 de um quadro inicial é automaticamente interpretado em *SPAD* como valor omissos, pois aquele assume como tabela condensada qualquer quadro onde conste algum valor diferente de 0 e de 1. Atendendo a estas e outras situações entretanto surgidas, foi necessário proceder à recodificação de praticamente todas as variáveis.

Logo após a primeira saída de resultados de uma AFCM do *SPAD*, demos conta de que este programa fornece automaticamente um procedimento que, perante as modalidades de menor peso, permite a reconstituição robusta dos resultados da análise. Esta operação é designada por “ventilação” e ocorre antes da determinação dos eixos de análise. Para cada questão, os indivíduos associados a uma modalidade considerada de peso fraco (peso inferior a 2% do peso total dos indivíduos¹⁴) serão redistribuídos ao acaso pelas restantes modalidades da questão a que pertencem. Desta forma, o quadro inicial conserva o seu carácter de “disjuntivo completo”, sendo o total de respostas a cada questão sempre igual a n (número de indivíduos inquiridos).

¹⁴ por convenção do *software*, podendo no entanto ser alterado pelo utilizador

Saliente-se que durante o estudo univariado fomos dando conta de algumas modalidades que estariam nestas condições e que classificamos como *outliers*. Poderíamos optar simplesmente por não as considerar, excluindo à partida os indivíduos associados a cada modalidade considerada de peso fraco. Mas, nesse caso estaríamos a perder informação, pondo de parte indivíduos da amostra que deixariam de poder dar a sua contribuição para a caracterização da população em estudo.

Com os valores omissos aconteceu algo semelhante. No estudo unidimensional observámos que também estes tinham, em geral, pouca expressão, considerando o conjunto das modalidades da questão em estudo. De facto, durante a organização dos dados verificamos que alguns alunos não responderam a uma ou outra questão. Poderíamos ter procedido à imputação parcial da resposta, usando um método adequado de tratamento de valores omissos. Pensamos não ser esta no entanto a melhor alternativa, uma vez que com este procedimento poderíamos perturbar a estrutura global das variáveis em análise. Esta primeira fase da análise factorial das correspondências múltiplas do SPAD deu-nos a confirmação que grande parte dos valores omissos não eram relevantes para o estudo, tornando-se “ventilados” na fase de apuramento.

Assumindo o quadro inicial como quadro disjuntivo completo, surgiu ainda um outro obstáculo associado às questões de escolha múltipla, onde os inquiridos deveriam escolher 4 entre 8 modalidades – nas questões III e IV – e 3 entre 6 modalidades – na questão IX. Por forma a tornar possível o tratamento destas questões em AFCM, cada modalidade passou a ser tomada como uma questão constituída por duas modalidades: “sim” ou “não”.

No que respeita à idade dos alunos, atendendo a que houve apenas 8 respostas distintas, considerámos esta variável como nominal, para poder ser incluída na análise multivariada. Como foi dito atrás, a profissão dos pais foi codificada a partir da “Classificação Internacional de Profissões (versão portuguesa)”, usando 3 algarismos¹⁵. Tendo resultado desta classificação um número elevado de modalidades de resposta, passamos a considerar 1 só algarismo para esta análise. As habilitações literárias dos pais foram agrupadas em 4 modalidades, tal como foi considerado no estudo bivariado.

Assim, do apuramento resultaram 149 modalidades e 49 questões¹⁶. Note-se que 28 das modalidades foram “ventiladas” nesta fase por terem um peso pouco significativo, muitas das quais correspondiam a valores omissos. O *software* edita o quadro de Burt, sendo neste caso uma tabela de ordem (149, 149).

O *output* da AFCM fornece ainda o histograma dos 100 primeiros valores próprios e a percentagem de inércia explicada por cada um dos eixos principais. Determinando o número de eixos a reter usando o critério de selecção que considera apenas os valores próprios superiores a $\frac{1}{T}$ (sendo T o número de questões) teremos

¹⁵ Esta Classificação foi extraída do livro “Antecedentes, Metodologias e Conceito – Censo 91”

¹⁶ O número de questões aumentou consideravelmente em relação às análises das secções anteriores, por termos assumido cada modalidade resultante das 3 questões de escolha múltipla como uma nova questão (em vez de 3 questões passamos a ter 24)

neste caso de reter os 37 primeiros eixos, correspondendo a cerca de 70% da inércia explicada.

Partindo do valor das coordenadas de cada uma das modalidades em cada eixo principal de inércia e do valor da respectiva contribuição absoluta, isto é, a contribuição de cada modalidade para a formação de cada eixo, podemos assim concluir quais os grupos de modalidades que estão associadas e por outro lado identificar grupos de variáveis que se opõem em cada eixo.

O SPAD fornece ainda a selecção das modalidades com mais significado em cada extremidade dos eixos, de extrema importância para a descrição dos factores de análise. A interpretação dos factores principais fica assim facilitada por este procedimento que fornece, não só os grupos de modalidades mais significantes em cada eixo, como também o valor das respectivas coordenadas e contribuições.

Caso o utilizador não faça outra opção, o programa parte de um critério estatístico designado por valor-teste para a classificação das modalidades, e não o das coordenadas. A ordenação dos elementos resultantes de um quadro complexo por valores-testes decrescentes faz realçar os aspectos mais salientes numa primeira análise, mostrando rapidamente as coerências e os elementos mais contrastantes, quer entre as modalidades quer entre os indivíduos ou mesmo entre as variáveis. O método está relacionado com a realização de testes estatísticos que permitem avaliar as diferenças entre proporções ou entre médias. Comparando, entretanto, a descrição dos factores principais segundo as coordenadas das modalidades nos eixos principais com a que parte dos respectivos valores-testes, optamos por esta última, pois o método das coordenadas assume todas as modalidades do questionário, ainda que tenham sido "ventiladas" depois do apuramento inicial, o que pensamos não ser o melhor caminho para esta análise.

4.3 INTERPRETAÇÃO DOS EIXOS PRINCIPAIS

O primeiro eixo principal de inércia (Gráfico 4.1) agrupa no extremo negativo os alunos com mau rendimento em Matemática durante todo o 3º ciclo (nível 2 nos 3 anos de escolaridade), que consideram que a Matemática não é importante para a sua formação e que optaram pelo 4º agrupamento de estudos no ensino secundário (Humanidades). No extremo oposto coloca os alunos de excelente rendimento no 3º ciclo (nível 5 nos 3 anos de escolaridade), que vêem na Matemática uma componente importante para a sua formação e que optaram pelo 1º agrupamento de estudos (Científica e Natural). Estes dois grupos distinguem-se ainda pelo facto dos segundos terem a Matemática como disciplina base do seu curso, isto é, como disciplina específica da componente de formação, enquanto os primeiros têm Métodos Quantitativos (apenas no 10º ou 11º ano de escolaridade) como disciplina técnica.

Saliente-se o facto de a representação gráfica, por si só, não ser conclusiva no que diz respeito ao posicionamento que ocupa uma dada modalidade num dado eixo. É necessário conhecer antes de mais a contribuição relativa da modalidade nesse eixo, isto é, saber-se se essa modalidade está ou não bem representada nesse eixo. No entanto, se a análise for realizada partindo dos valores da contribuição das modalidades para a formação de cada eixo principal (contribuições absolutas), pode prescindir-se da contribuição relativa.

No segundo eixo podemos ver associadas as variáveis: rendimento em Matemática; capacidades dos professores desta disciplina; habilitações literárias e profissões dos pais dos alunos; e ainda a opção por cursos de carácter geral ou cursos tecnológicos. Por um lado, temos alunos que tiveram o mesmo professor de Matemática ao longo do 3º ciclo, professor este que os alunos consideram ser o melhor possível, com muito boas capacidades científicas, pedagógicas e de relacionamento com os alunos. Estes alunos obtiveram sempre nível 5 durante o 3º ciclo e são filhos de pais com profissões intelectuais e científicas (PIC) cujas habilitações literárias são, no mínimo, licenciaturas. No outro extremo temos alunos dos cursos tecnológicos cujos pais não têm mais que o 1º ciclo e, não tendo tido o mesmo professor nos três anos de escolaridade, consideram que praticamente todos os seus professores de Matemática foram razoáveis, quer em termos de capacidades científicas e pedagógicas quer em termos de capacidades de relacionamento com os seus alunos. Estes alunos atingiram pouco mais do que os objectivos mínimos durante o 3º ciclo, pois obtiveram nível 3 em todos os anos de escolaridade deste ciclo.

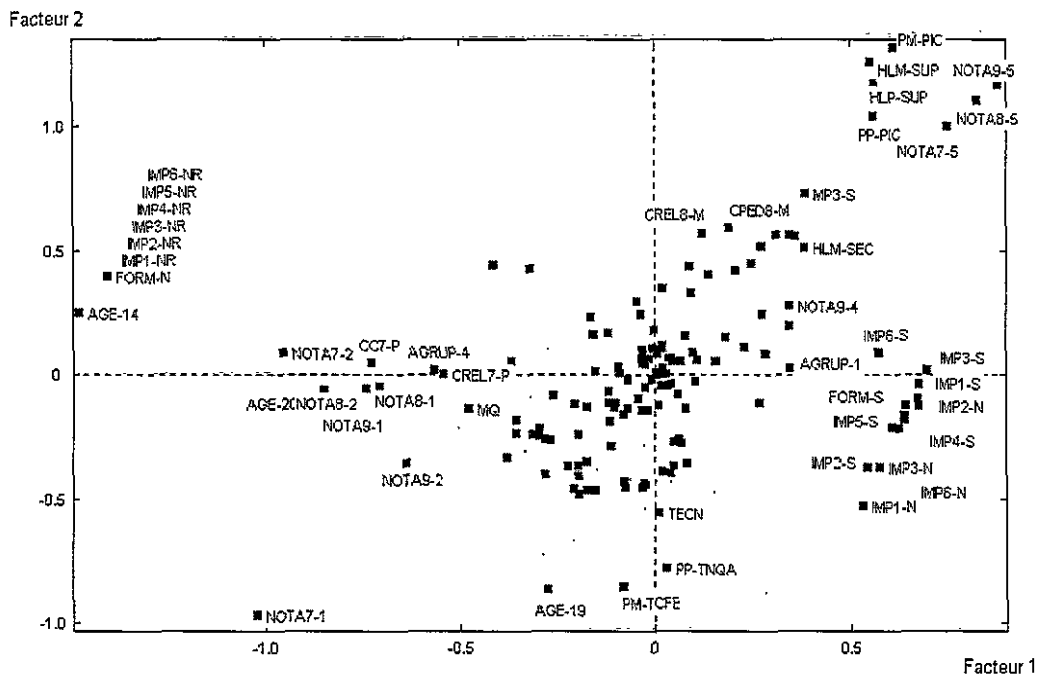


Gráfico 4.1 Representação das modalidades no plano formado pelo 1º e 2º factores principais

Por sua vez, o terceiro eixo (Gráfico 4.2) põe em oposição este último grupo de alunos com os que afirmam terem tido fracos professores de Matemática, pois apresentavam poucas capacidades de relacionamento, pedagógicas e científicas. Este segundo grupo de alunos tem a particularidade de ter omitido no inquérito os seus resultados no que respeita às classificações em Matemática.

Um grupo de alunos do 2º agrupamento (Artes) aparece evidenciado no quarto eixo, caracterizando-se por terem tido o mesmo professor no 8º e 9º anos. Este professor é considerado (pelos alunos) muito bom, em termos pedagógicos e de relacionamento. Em oposição temos um outro grupo que se salienta pelo facto de terem tido um professor diferente em cada ano, em que as capacidades científicas,

pedagógicas e de relacionamento dos professores dos dois últimos anos foram consideradas (segundo os alunos) muito fracas. O primeiro grupo de estudantes justifica o insucesso em Matemática essencialmente devido à falta de empenho dos alunos, enquanto os seus opositores neste eixo atribuem aos manuais escolares pouco interessantes a causa desta situação.

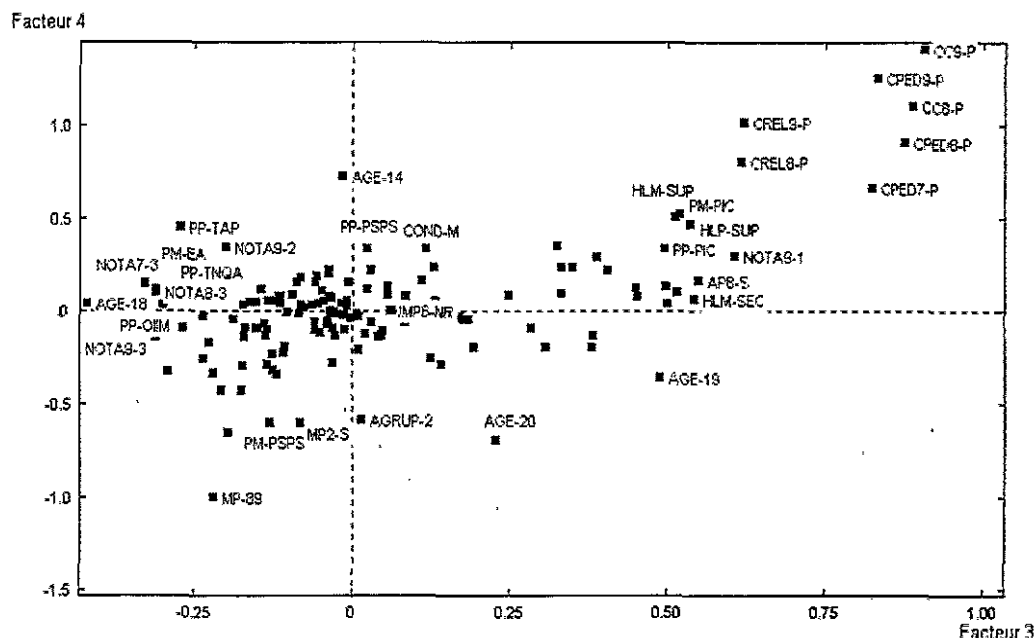


Gráfico 4.2 Representação das modalidades no plano formado pelo 3º e 4º factores principais

O 5º eixo factorial põe em oposição essencialmente os alunos que tiveram o mesmo professor em dois anos com os que tiveram sempre professores diferentes. Os primeiros são alunos de Humanidades, tendo assim Métodos Quantitativos como disciplina técnica, enquanto os seus opositores têm Matemática como disciplina específica. Em relação aos factores de maior peso para a aprendizagem da Matemática, os que experimentaram dois anos consecutivos com o mesmo professor pensam ser o interesse prático dos temas e a adopção de um bom manual. Os outros acham que o gosto pela Matemática é essencial para a aprendizagem desta disciplina.

No 6º eixo temos sobretudo alunos que tiveram professores com poucos atributos em termos de docência, opondo-se aos de razoáveis capacidades científicas, pedagógicas e de relacionamento. Os primeiros consideram que o que mais contribui para a aprendizagem da Matemática é o interesse prático dos temas, enquanto os que se lhe opõem neste eixo consideram não ser este o factor mais importante.

Estudantes do sexo masculino encontram-se em oposição com um grupo de alunas, no 7º eixo factorial. Os rapazes optam por agrupamentos em que Matemática é uma disciplina específica. As raparigas abandonam mais cedo a Matemática, escolhendo o curso de Humanidades que tem apenas mais um ano com Métodos Quantitativos. Neste eixo podemos ainda observar que os rapazes classificam de razoáveis os seus professores de Matemática (em relação aos três parâmetros em estudo). Em contra partida, as raparigas consideram que encontraram durante o 3º ciclo sempre muito bons professores. Curiosamente estas alunas têm pais que frequentaram no máximo até ao 4º ano do 1º ciclo, sendo a maioria das mães empregadas administrativas (EA) e os pais trabalhadores da produção industrial ou

artesãos (TPIA). Quer os pais quer as mães dos rapazes têm habilitações literárias superiores e exercem profissões designadas por intelectuais e científicas (PIC). Podemos ainda verificar que este grupo de rapazes defende que a adopção de um bom manual é um factor que muito contribui para a aprendizagem da Matemática. As raparigas, pelo contrário, pensam que o gosto pela Matemática é que é mais determinante neste processo.

O 8º eixo é, por excelência, o eixo dos grupos profissionais dos pais dos alunos inquiridos e respectivas habilitações literárias. No extremo negativo temos pais cujas habilitações literárias são apenas o 1º ciclo (completo ou incompleto) ou habilitações superiores. No extremo oposto vemos habilitações mais intermédias - 2º, 3º ciclo e secundário. Do lado esquerdo deste eixo temos pais artesãos e trabalhadores da produção industrial (TPIA) e ainda os que desempenham profissões intelectuais e científica (PIC). À direita temos alunos cujos pais são técnicos das ciências físicas e de engenharia (TCFE); pessoal dos serviços de protecção e segurança, dos serviços pessoais e domésticos e trabalhadores similares (PSPS). Ainda no extremo direito deste eixo, vemos que as mães destes alunos são em geral trabalhadoras não qualificadas da agricultura, indústria, comércio e serviços (TNQA); trabalhadoras da agricultura e da pesca (TAP); encontrando-se ainda muito bem representadas as mães que ocupam lugares de direcção em empresas (DQD). É importante salientar que os alunos caracterizados no extremo direito do 8º eixo, durante o 3º ciclo, foram sempre bons alunos atingindo o nível 4 todos os anos. No outro extremo temos alunos com nível 2 e outros com nível 5 nos três anos de escolaridade.

O 9º eixo factorial agrupa alunos cujos pais frequentaram o 2º ou o 3º ciclo opondo-se aos que têm frequência ao nível do secundário. Os primeiros pais são trabalhadores da agricultura ou da pesca - com ou sem qualificação (TAP e TNQA), enquanto os últimos são empregados administrativos (EA) ou directores de empresas ou da função pública (DQD). Os dois grupos distinguem-se ainda pelo facto dos alunos cujos pais têm menos habilitações literárias considerarem que o empenhamento do aluno é um factor com grande contribuição para a aprendizagem da Matemática. Os alunos que neste eixo se lhes opõem pensam ser o interesse prático dos temas que mais influencia a aprendizagem nesta disciplina. Ambos os grupos vêem na Matemática uma componente importante na sua formação, embora por diferentes razões: os primeiros porque dizem ser a base para a compreensão de outras ciências, os outros consideram que é o facto da Matemática facilitar a interpretação da informação quantitativa divulgada pelos meios de comunicação.

No 10º eixo temos de novo a oposição entre os rapazes do 1º agrupamento e as raparigas do 4º. Eles tiveram sempre o mesmo professor de Matemática durante o 3º ciclo e elas apenas no 7º e no 8º ano. Diferem também na selecção dos factores que mais contribuem para o insucesso da Matemática: os rapazes dizem dever-se à natureza abstracta da disciplina; as raparigas atribuem à sua falta de conhecimentos básicos, assim como à falta de empenho. Podemos verificar ainda que enquanto as raparigas defendem que o gosto pela Matemática, o empenhamento do aluno e a competência científica e pedagógica dos professores são os factores mais importantes para a aprendizagem da Matemática, os rapazes consideram ser mais determinante a utilização de diferentes materiais didácticos na aula e a adopção de um bom manual escolar.

No 11º eixo temos, por um lado, alunos do 1º agrupamento cujo rendimento em Matemática foi muito bom durante o 3º ciclo e que vivem com os pais; por outro, temos alunos do 4º agrupamento que vivem apenas com o pai ou a mãe. Os primeiros são de opinião que a Matemática é importante para a sua formação, por ser a base para a compreensão de outras ciências. Evidencia-se neste eixo o facto dos alunos de Humanidades reconhecerem a importância da Matemática como uma componente da sua formação, embora pelo facto desta facilitar a interpretação da informação divulgada pelos meios de comunicação (informação esta de resto já evidenciada no 9º eixo).

Verificamos no 12º eixo a oposição entre os alunos que consideram que o seu aproveitamento em Matemática em nada condicionou a opção do agrupamento de estudo para o secundário (alunos que efectivamente escolheram Humanidades - agrupamento 4), e os que afirmam ter tido algum peso (alunos que optaram pelo curso de Científica e Natural - agrupamento 1). De notar uma outra característica da vivência escolar (no que respeita à Matemática) destes alunos: os do 1º agrupamento tiveram sempre o mesmo professor, o mesmo não acontecendo com os do 4º. Estes 2 grupos aparecem neste eixo em extremos opostos também por terem opiniões diferentes no que respeita aos factores de insucesso na disciplina de Matemática: os alunos de Humanidades atribuem responsabilidades aos manuais escolares pouco interessantes e à abstracção da disciplina; os do 1º agrupamento dizem dever-se à falta de empenho e de conhecimentos básicos dos alunos. Curiosamente, estas opiniões são exactamente contrárias às que verificamos no 9º eixo factorial (em relação a alunos pertencentes aos mesmos agrupamentos).

Alunos do 3º agrupamento (Económico-Social), que tiveram sempre professores diferentes, aparecem em oposição no 13º eixo com os do 1º que tiveram o mesmo professor em pelo menos dois anos do 3º ciclo. O contraste entre eles explica-se também pelas diferentes razões evocadas para a aprendizagem e insucesso da Matemática: a adopção de um bom manual escolar e a competência pedagógica e científica dos professores, são os aspectos apresentados pelos primeiros alunos em relação à aprendizagem; os alunos do 1º agrupamento consideram fundamental a utilização de diferentes materiais didácticos na aula. O insucesso em Matemática, segundo os alunos do agrupamento 3 representados neste eixo, deve-se à grande abstracção da disciplina, enquanto este aspecto não é considerado relevante para os seus opositores neste eixo.

O 14º eixo caracteriza dois grupos de alunos que se distinguem essencialmente pela influência que o aproveitamento em Matemática teve na decisão que tomaram em termos de agrupamento de estudos. No extremo negativo temos os que dizem que o seu rendimento em nada condicionou a decisão final, e no positivo os que se sentiram muito condicionados. Estes últimos, por sinal, tiveram sempre o mesmo professor, o que não aconteceu com os outros. As habilitações literárias dos pais dos alunos destes dois grupos é outra razão que os opõe - à esquerda temos alunos cujos pais têm no máximo o 4º ano e à direita os que têm entre 5 e 12 anos de escolaridade.

Acumulando a informação até ao 14º eixo, temos cerca de 42% da inércia total. A informação mais relevante é sem dúvida a fornecida pelos primeiros eixos. As modalidades em oposição nos últimos eixos descritos não são de todo as mais significativas neste estudo. Assim, não julgamos útil prolongar a análise, já que atingimos uma zona onde a informação remanescente é praticamente repartida em partes iguais pelos eixos seguintes, retratando comportamentos particulares observados e não tanto associações relevantes.

4.4 CONCLUSÕES

Acreditamos ser distante o tempo em que nas Escolas a vida tinha um carácter rotineiro (salvo honrosas excepções!), em que tudo o que havia para aprender estava nos livros e registado na memória do professor, enquanto aos alunos era exigida uma atitude passiva, sendo valorizada pouco mais que a memorização. Apesar das fortes críticas de que foi alvo este sistema, de certa forma era coerente com as exigências do mundo do trabalho. A preparação dos jovens para o mundo do trabalho era feita com base numa actividade que muito provavelmente se iria manter durante toda a sua vida, sem grandes alterações dos gestos profissionais exigidos.

Hoje, em virtude das grandes transformações tecnológicas, mais de que nunca, os conhecimentos utilizáveis no plano profissional são rapidamente desactualizados. Sabemos que este vertiginoso progresso, capaz de produzir cada vez mais bens e serviços, exige cada vez mais uma grande capacidade de adaptação a novas situações, sendo fundamental desenvolver um maior número de competências.

É de facto urgente que a Escola se dê conta do papel que tem a desempenhar. A Escola terá de deixar de se restringir à transmissão de conhecimentos e impulsionar o alargamento da formação dos jovens de hoje a áreas como a aquisição de competências de raciocínio, de processamento e recolha de informação, de relacionamento e de investigação. A este propósito afirma Imaginário (1987, p. 57), que a Escola deve preocupar-se com “o saber, o saber-fazer, o saber-ser e o saber-tornar-se”.

Com o questionário que aplicámos a uma amostra de alunos do 10º ano de escolaridade, das Escolas Públicas do Distrito do Porto, pretendíamos encontrar associações entre o rendimento do aluno em Matemática e um conjunto de parâmetros, uns que dizem respeito ao contexto familiar do aluno (a actividade profissional e o grau de instrução dos respectivos pais, e parentesco com as pessoas com quem o aluno vive) e outros que se relacionam com as características dos professores de Matemática que tiveram - nomeadamente nos três anos que antecedem o ensino secundário.

Verificamos existir de facto uma associação positiva entre o aproveitamento dos alunos em Matemática no 3º ciclo e as habilitações literárias dos respectivos pais. No entanto, o grau de associação encontrado aponta para uma associação fraca. O mesmo acontece em relação às competências do professor de Matemática (avaliadas pelo aluno) e o sucesso do aluno: rejeitámos com firmeza a hipótese de independência entre estas variáveis, mas os resultados encontrados indicam que as associações existentes, embora todas no sentido positivo, são fracas. Podemos assim inferir que o desempenho dos alunos em Matemática está condicionado por factores como a capacidade científica, pedagógica e de relacionamento do professor desta disciplina, no entanto, estes factores, por si só, não são determinantes do bom rendimento do aluno nem explicam o seu eventual insucesso.

Procurando estudar as razões evocadas pelos alunos, que mais fortemente influenciam a sua aprendizagem da Matemática, verificamos que 85% dos alunos consideram ser o “empenhamento do aluno” e cerca de 80% seleccionam o “gosto pela Matemática” e a “competência pedagógica e científica do professor”.

O motivo mais apontado pelos alunos para o insucesso da disciplina em causa foi “a falta de conhecimentos básicos” dos alunos (84%). Curiosamente, 75% dos alunos inquiridos consideram que a “falta de empenho dos alunos” é um factor que muito contribui para esta situação. Destacamos ainda “a monotonia das aulas” que para muitos é também geradora de mau rendimento na disciplina.

Salientamos o facto de 30% dos alunos considerarem que a Matemática não tem nenhum interesse para a sua formação. Por outro lado, cerca de 75% dos alunos afirmam que a sua opção da área de estudos, para o prosseguimento de estudos no ensino secundário, ficou de algum modo condicionado pelo rendimento obtido na disciplina de Matemática.

As razões que explicam a importância atribuída pelos alunos à disciplina de Matemática são: a sua aplicação em muitas actividades profissionais (77%) e o facto de permitir o desenvolvimento da capacidade de raciocínio e de abstracção (72%).

A análise multivariada levada a cabo estabelece uma interessante hierarquia quanto aos antecedentes que influenciaram directamente ou indirectamente as opções de estudo no ensino secundário. Destaca-se em primeiro lugar o facto de o nível de rendimento escolar e especialmente a dicotomia entre insucesso e sucesso produzir um efeito na escolha dos agrupamentos que incluam ou não a Matemática como disciplina específica.

Em segundo lugar, o estudo revela a conotação positiva do enquadramento social do aluno avaliado pelas habilitações literárias dos pais e respectivas profissões, na taxa de sucesso escolar em Matemática.

Por fim, a análise multivariada evidencia o peso negativo que os professores considerados fracos pelos alunos exerce na adesão dos mesmos a este tipo de matéria condicionando o seu sucesso escolar e as opções fundamentais em termos de formação profissional.

O trabalho subjacente a este estudo procurou contribuir para a compreensão das dificuldades na disciplina de Matemática e para a discussão em torno das estratégias para a sua atenuação, abordando a questão do insucesso em Matemática na perspectiva do aluno e as implicações que o mau rendimento pode ter na decisão a tomar em termos de escolha da área de estudos a seguir e futuras opções profissionais.

REFERÊNCIAS BIBLIOGRÁFICAS

- CRAMER, D. (1998). *Fundamental Statistics for Social Research. Step-by-step calculations and computer techniques using SPSS for Windows*. London: Routledge.
- GHIGLIONE, R. & MATALON, B. (1993). *O Inquérito – Teoria e Prática*. Oeiras: Celta Editora Lda.
- GOMES, P. (1993). *Análise de Dados*. Lisboa: Instituto Superior de Estatística e Gestão de Informação. Universidade Nova de Lisboa.
- GOMES, P. (1998). *Tópicos de Sondagens*. Lisboa: Sociedade Portuguesa de Estatística.
- GUIMARÃES, R. C. & CABRAL, J. S. (1997). *Estatística*. Alfragide: Editora McGraw-Hill de Portugal, Lda.
- MURTEIRA, B. J. F. (1993). *Análise Exploratória de Dados – Estatística Descritiva*. Lisboa: McGraw-Hill de Portugal, Lda.
- MURTEIRA, B. J. F. (1990). *Probabilidades e Estatística Vol. I e II*. Lisboa: McGraw-Hill de Portugal, Lda.
- NCTM (1994). *Normas Profissionais para o Ensino da Matemática*. Lisboa: APM e IIE.
- NCTM (1991). *Normas para o Currículo e a Avaliação em Matemática Escolar*. Lisboa: APM e IIE.
- REIS, E. (1997). *Estatística Multivariada Aplicada*. Lisboa: Edições Sílabo.
- VICENTE, P. & REIS, E. & FERRÃO, F. (1996). *Sondagens – A amostragem como factor decisivo da qualidade*. Lisboa: Edições Sílabo, Lda.

ANEXO - Questionário aplicado aos alunos

1. Identificação

Escola Turma

Agrupamento Curso: Carácter geral Tecnológico. Qual?

Idade: anos Sexo: Masculino Feminino

Habilitações literárias: Pai Mãe

Profissão: Pai Mãe

Com quem vive?

NOTE: Se não vive com os seus pais, indique, relativamente à pessoa mais próxima que vive consigo: habilitações literárias: profissão:

2. A disciplina de Matemática no percurso do 3º ciclo (7º, 8º e 9º anos)

ASSINALE COM UMA CRUZ (X) A SUA RESPOSTA A CADA UMA DAS QUESTÕES SEGUINTE

NOTE: Se repetiu algum destes três anos, responda tendo em consideração apenas o ano de passagem

I. Teve o mesmo professor nos três anos?

Sim	Não
<input type="checkbox"/>	<input type="checkbox"/>
<input type="checkbox"/>	<input type="checkbox"/>

Teve o mesmo professor apenas em dois desses anos?

Se respondeu "sim", indique esses dois anos? __, __

II. Como caracteriza os seus professores de Matemática, em termos de:

	7º Ano			8º Ano			9º Ano		
	Pouca	Razoável	Muita	Pouca	Razoável	Muita	Pouca	Razoável	Muita
Capacidade de relacionamento	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Competência pedagógica	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Competência científica	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

III. Indique **quatro** factores que, na sua opinião, mais contribuem para a **aprendizagem** da disciplina de Matemática:

- Gosto pela Matemática
- Adopção de um bom manual escolar
- Empenhamento do aluno
- Utilização de diferentes materiais didácticos na aula
- Capacidade intelectual do aluno
- Competência científica e pedagógica do professor
- Interesse prático dos temas
- Outros

IV. Indique quatro factores que, na sua opinião, mais contribuem para o insucesso da disciplina de Matemática:

- | | |
|---|--|
| <input type="checkbox"/> Falta de conhecimentos básicos | <input type="checkbox"/> Professores demasiado exigentes |
| <input type="checkbox"/> Desinteresse dos temas | <input type="checkbox"/> Manuais escolares pouco interessantes |
| <input type="checkbox"/> Falta de empenho dos alunos | <input type="checkbox"/> Monotonia das aulas |
| <input type="checkbox"/> Disciplina muito abstracta | <input type="checkbox"/> Outros..... |

V. Quais os níveis que obteve no 3º período em cada um dos anos do 3º ciclo?

	7º Ano	8º Ano	9º Ano
1			
2			
3			
4			
5			

VI. No 10º ano tem a disciplina de: Matemática Métodos Quantitativos

VII. O seu aproveitamento na disciplina de Matemática durante o 3º ciclo, condicionou a sua opção do agrupamento de estudos para o secundário?

Muito	Pouco	Nada

VIII. Considera a disciplina de Matemática importante para a sua formação? Sim Não

IX. Se respondeu "sim" à questão anterior, seleccione, entre as razões seguintes, as três que melhor explicam a importância que atribui à disciplina de Matemática:

- | | |
|--|---|
| <input type="checkbox"/> Desenvolve a capacidade de raciocínio e abstracção. | <input type="checkbox"/> Tem aplicação em muitas actividades Profissionais. |
| <input type="checkbox"/> Facilita a interpretação da informação quantitativa divulgada pelos meios de comunicação. | <input type="checkbox"/> Está presente em diversas situações do nosso quotidiano. |
| <input type="checkbox"/> É a base para a compreensão de outras ciências. | <input type="checkbox"/> Outros. |

**APPLICATION OF THE MODIFIED "PEAK" TO
CHARACTERIZE THE $M | G | \infty$ QUEUE BUSY CYCLE
LENGTH DISTRIBUTION**

Autor:
Manuel Alberto Martins Ferreira
UNIDE/ISCTE

APPLICATION OF THE MODIFIED "PEAK" TO CHARACTERIZE
THE $M|G|\infty$ QUEUE BUSY CYCLE LENGTH DISTRIBUTION

APLICAÇÃO DO "PICO" MODIFICADO NA CARACTERIZAÇÃO
DA DISTRIBUIÇÃO DO COMPRIMENTO DO CICLO DE
OCUPAÇÃO DA FILA DE ESPERA $M|G|\infty$

Autor: Manuel Alberto Martins Ferreira
- UNIDE/ISCTE

Abstract

- We present a parameter to characterise the $M|G|\infty$ queue busy cycle length distribution that is a modification of the "peak" presented by (Whitt, 1984). We already did a similar work with the $M|G|\infty$ queue busy period length distribution (Ferreira, 1997).

KEY WORDS:

- $M|G|\infty$, busy cycle, "peak", parameter q .

Resumo

- Neste trabalho propomos um parâmetro para caracterizar a distribuição do comprimento do ciclo de ocupação da fila de espera $M|G|\infty$ que é uma modificação do "pico" proposto por (Whitt, 1984). Pode ver-se um estudo análogo para o comprimento do período de ocupação da fila de espera $M|G|\infty$ em (Ferreira, 1997).

PALAVRAS-CHAVE:

- $M|G|\infty$, ciclo de ocupação, "pico", parâmetro q .

1. INTRODUCTION

The busy cycle length of the MIG_{∞} queue is a random variable, Z , with Laplace transform (Takács, 1962)

$$\bar{Z}(s) = \bar{I}(s) \bar{B}(s) \quad (1.1)$$

where

$$\bar{I}(s) = \frac{\lambda}{\lambda + s} \quad (1.2)$$

and

$$\bar{B}(s) = 1 + \frac{1}{\lambda} \left(s - \frac{1}{\int_0^{\infty} e^{-st - \lambda \int_0^t [1-G(v)] dv} dt} \right) \quad (1.3)$$

being λ the Poisson process arrival rate and $G(\cdot)$ the service distribution function.

Its inversion is a complicate task except for the collection of service time distributions (Ferreira, 1998)

$$G(t) = 1 - \frac{(1 - e^{-\rho})(\lambda + \beta)}{\lambda e^{-\rho} (e^{(\lambda + \beta)t} - 1) + \lambda}, \quad t \geq 0, \quad -\lambda \leq \beta \leq \frac{\lambda}{e^{\rho} - 1} \quad (1.4)$$

($\rho = \lambda\alpha$ being α the mean service time). In this case the distribution of Z is a mixture of exponentials:

$$P(Z \leq t) = 1 - \frac{(1 - e^{-\rho})(\lambda + \beta)}{\lambda - e^{-\rho}(\lambda + \beta)} e^{-e^{-\rho}(\lambda + \beta)t} + \frac{\beta}{\lambda - e^{-\rho}(\lambda + \beta)} e^{-\lambda t},$$

$$t \geq 0, -\lambda \leq \beta \leq \frac{\lambda}{e^{\rho} - 1}$$

(1.5)

To the $MID|_{\infty}$ system (constant service with value α) it is possible to invert (1.1) but we get a complicate expression, not useful to compute probabilities, although interpretable.

We can also get expressions for the moments of Z . They are complicate and demand numerical methods with the exception of

- Service times given by (1.4),
- Service times constant,
- First order moment for any service time distribution:

$$E[Z] = \frac{e^{\rho}}{\lambda} \quad (1.6)$$

due to (1.3), of course.

So we have an easy situation to the $MIG|_{\infty}$ systems with service time distribution functions given by (1.4). To the $MID|_{\infty}$ systems it is good only for the moments. For other systems, namely to the $MIM|_{\infty}$ systems (exponential service time) we have complicate situations.

Then we will present a parameter to characterise $MIG|_{\infty}$ queue system busy cycle length distribution. It is a modification of the "peak" proposed by Whitt (1984). Its expression, for various service time distributions, namely for the exponential distribution, is very simple.

THE "PEAK"

(Whitt, 1984) propose the use of the "peak", that is the value of the Laplace transform at $\frac{1}{\alpha}$, to characterise probability distributions related to queues, instead of the variance.

For the busy cycle length of the MIG_1^∞ queue, calling p to the "peak", we have from (1.1)

$$p = \bar{Z}\left(\frac{1}{\alpha}\right) = 1 - \frac{\alpha}{(\rho - 1) \int_0^\infty e^{-\frac{t}{\alpha} - \lambda \int_0^t [1-G(v)]dv} dt} \quad (2.1)$$

and

$$p = \sum_{n=0}^{\infty} (-1)^n \frac{E[Z^n]}{n! \alpha^n} \quad (2.2)$$

So p incorporates information of the moments of every order of Z .

To the collection of service time distributions given by (1.4):

$$p = \alpha \frac{e^{-\rho} (\lambda + \beta)(\rho + 1) - \beta}{(\rho + 1)(e^{-\rho}(\rho + \alpha\beta) + 1)}, \quad -\lambda \leq \beta \leq \frac{\lambda}{e^\rho - 1} \quad (2.3)$$

and

$$- \beta = -\lambda$$

$$p = 0 \quad (2.4)$$

$$- \beta = 0 \text{ (MIG}_1^\infty \text{ system)}$$

$$p^{G_1} = \frac{\rho}{e^\rho + \rho} \quad (2.5)$$

- $\beta = \frac{\lambda}{e^\rho - 1}$ (MIG₂| ∞ system)

$$p^{G_2} = \frac{\rho^2}{e^\rho(\rho+1) + \rho^2 - 1} \quad (2.6)$$

For the systems

- MID| ∞

$$p^D = \frac{\rho}{e^{\rho+1} + \rho} \quad (2.7)$$

- MIM| ∞

$$p^M = \frac{e^\rho - \rho - 1}{(\rho+1)(e^\rho - 1)} \quad (2.8)$$

In G_1, G_2, D and M cases p depends only on ρ through very simple expressions. Except $E[Z]$ it is the only parameter with simple expression, for the distribution of Z , to the MIM| ∞ systems.

To service time distributions related with the exponential one we still have:

- If the service time distribution is NBUE (new better than used in expectation)

with mean α , $\int_t^\infty [1 - G(v)]dv \leq \int_t^\infty e^{-\frac{v}{\alpha}}dv$ (Ross, 1983, page 273) and, so,

$$p^{NBUE} \leq \frac{e^\rho - \rho - 1}{(\rho + 1)(e^\rho - 1)} \quad (2.9)$$

- If the service time distribution is NWUE (new worse than used in expectation)

with mean α , $\int_t^\infty [1 - G(v)]dv \geq \int_t^\infty e^{-\frac{v}{\alpha}} dv$ (Ross, 1983, page 273) and, so,

$$p^{NWUE} \geq \frac{e^\rho - \rho - 1}{(\rho + 1)(e^\rho - 1)} \quad (2.10)$$

And if $G(\cdot)$ and $H(\cdot)$ are service time distributions such that $1 - G(t) \leq 1 - H(t)$

$$p^G \geq p^H \quad (2.11)$$

2. THE PARAMETER Q

The expression (2.2) can be written as $p = 1 - \frac{e^\rho}{\rho} + \sum_{n=2}^{\infty} (-1)^n \frac{E[Z^n]}{n! \alpha^n}$. As $1 - \frac{e^\rho}{\rho}$

does not depend on the service time distribution we put

$$q = p \frac{\rho}{e^\rho - \rho} + 1 \quad (3.1)$$

We take out of p the parts that do not depend on the service time distributions form. Then we normalise dividing for the common part.

Of course we can write simple expressions for q^{G_1} , q^{G_2} , q^D and q^M depending only on ρ . And, evidently, q has the same properties of p .

It is easy to show that

$$1 \leq q \leq \frac{\rho^2}{(e^\rho + 1)(e^\rho - \rho)} + 1 \quad (3.2)$$

for any service time distribution. So

$$\lim_{\rho \rightarrow \infty} q = 1 \quad (3.3)$$

We compute q for some service time distributions, taking various values of ρ and we got

G	ρ	.5	1	5	10	15	20
G ₁		1.0573588	1.0841882	1.0000456	1.0000000	1.0000000	1.0000000
G ₂		1.0357639	1.0575796	1.0000383	1.0000000	1.0000000	1.0000000
D		1.0247402	1.0373147	1.0000171	1.0000000	1.0000000	1.0000000
M		1.0376733	1.0654280	1.0002254	1.0000000	1.0000000	1.0000000
P		1.0359789	1.0646113	1.0004005	1.0000001	1.0000000	1.0000000

being P the power distribution with parameter $c = .5 \left(\alpha = \frac{c}{c+1} \right)$. For this case q^P is computed directly from (2.1) and (3.1).

For great values of ρ (above $\rho = 10$ for some service time distributions and above $\rho = 20$ for others) under some conditions the busy period is practically exponentially distributed (Ferreira e Ramalhoto, 1994) and (Ferreira, 1996). In the same conditions q is practically 1.

3. CONCLUDING REMARKS

The parameter q defined in (3.1) incorporates information of the moments of every order of Z .

It takes different values when the distributions of Z change with the service time distribution.

There are simple bounds for q depending only on ρ (expression (3.2)).

Even for the $M|M|∞$ queue systems the expression of q is very simple.

It is admissible to conjecture that the same values of q , for different service time distributions, indicate similar behaviours for the corresponding distribution of Z .

So q may help to characterise the distribution of Z .

4. REFERENCES

- FERREIRA, M.A.M. e RAMALHOTO, M.F., (1994), "Estudo dos parâmetros básicos da fila de espera $M|G|∞$ ". *A Estatística e o Futuro e o Futuro da Estatística. Actas do I Congresso Anual de S.P.E.. Lisboa. 1994.*
- FERREIRA, M.A.M., (1996), "Distribuição do comprimento do período de ocupação da fila de espera $M|G|∞$ em situação de "Heavy Traffic". (Distribuições de serviço potência e de Pareto)". *Comunicação apresentada no IV Congresso Anual de S.P.E.. Funchal.*
- FERREIRA, M.A.M., (1997), "Um parâmetro caracterizador de Distribuição do Período de Ocupação da fila de espera $M|G|∞$ ". *Revista Portuguesa de Gestão. III/97. Lisboa.*
- FERREIRA, M.A.M., (1998), "Aplicação de equação de Ricatti ao estudo do período de ocupação do sistema $M|G|∞$ ". *Revista de Estatística. INE. Vol. 1. 1º quadrimestre.*
- ROSS, S., (1983), "Stochastic Processes". *Wiley. New York.*
- TAKÁCS, L., (1962), "An introduction to queueing theory". *Oxford University Press. New York.*
- WHITT, W., (1984), "On approximations for queues, I. Extremal distributions". *AT & T Laboratories Technical Journal, 63 (a). 115-138.*

FACTORES EXPLICATIVOS DO CONSUMO DE TABACO NUM DISTRITO DO INTERIOR DE PORTUGAL

Autor:
José Ramos Pires Manso

FACTORES EXPLICATIVOS DO CONSUMO DE TABACO NUM DISTRITO DO INTERIOR DE PORTUGAL

DETERMINANTS OF THE CONSUMPTION OF TOBACCO IN AN INTERIOR DISTRICT OF PORTUGAL

Autor: José Ramos Pires Manso, PhD
Professor Associado, Presidente da Comissão Científica
Do Departamento de Gestão e Economia
da Universidade da Beira Interior

RESUMO

- São objectivos deste 'working paper': (a) estudar os determinantes do consumo de tabaco nos jovens dos 12 aos 16 anos de um distrito do interior (norte) do país; (b) estudar a influência de factores como a idade, a escola, o absentismo às aulas, o insucesso escolar, o local de residência, o número de irmãos, a escolaridade dos pais, a situação conjugal destes, a religião, o número de horas de sono, o consumo de tabaco pelos pais, amigos e namorado(a), o consumo de bebidas alcoólicas e finalmente o efeito da depressão; (c) apreciar a coincidência, o paralelismo ou a obliquidade ou conflituabilidade dos comportamentos dos dois sexos face ao consumo de tabaco.

Palavras-chave:

- *bio-estatística, inferência estatística, factores comportamentais, comportamento da juventude, factores de consumo de tabaco*

ABSTRACT

- The main objectives of this working paper are: (a) to study the determinants of the consumption of tobacco by the young boys and girls of 12 to 16 years old in an (north) interior district of Portugal; (b) to study the influence of factors such as the age, the school, the absence to classes, the scholar insuccess, the place of residence, the number of brothers and sisters, the scholarship of their father and mother, and her marital situation, the religion, the number of sleeping hours, the consumption of tobacco by their fathers, friends and boy- or girl-friends, the consumption of alcoholic drinks and, finally, the depressive effect; (c) to appreciate the coincidence, the parallelism or the obliquity of the behaviour of both sexes as far as the consumption of tobacco concerns.

Key-words:

- *bio-statistics, statistical inference, behavioural factors, youth behaviour, tobacco consumption factors*

FACTORES EXPLICATIVOS DO CONSUMO DE TABACO NUM DISTRITO DO INTERIOR DE PORTUGAL

1. INTRODUÇÃO

No que se deixa em seguida vai ter-se em atenção o disposto por uma metodologia sugerida pela Organização Mundial de Saúde e por diversos autores, entre os quais, Ramalheira e Cardoso (1995). Essa metodologia vai ser aplicada aos dados de uma amostra relativamente grande de jovens de ambos os sexos dos 12 aos 16 anos que frequentam 6 escolas – repartidas em urbanas, suburbanas e rurais; com os elementos obtidos vai calcular-se algumas medidas típicas da estatística descritiva como a idade média (geral) de iniciação ao tabaco de cada um dos sexos, as frequências absolutas ou relativas (pesos) dos diversos atributos (escola, idade, ...) no total; numa segunda fase, calculam-se e interpretam-se algumas medidas de inferência estatística típicas de disciplinas como bioestatística: rácios de controlo, O_c , rácios de produtos cruzados, RC , rácios de risco, RR , e rácio de doentes, O_d

2. FRAMEWORK TEÓRICA

O que vamos ver vai seguir as seguintes etapas: Depois de arranjar um grupo de referência, a que chamaremos “grupo de controlo”, e de ver qual a incidência do fenómeno – consumo de tabaco – sobre esse grupo, isto é, de calcular o odds ou rácio de controlo, OR ou O_c , vai-se comparar este rácio ou odds do grupo com o rácio do grupo ‘doente’, isto é, o rácio odds do “grupo doentes”, O_d . O passo seguinte consiste na estimação do rácio de produtos cruzados, RC ou odds ratio, OR , rácio que é geralmente entendido como uma boa estimativa do rácio de risco, RR , que está associado a uma determinada doença.

Faz-se em seguida a interpretação dos resultados através da realização de um teste estatístico cuja hipótese nula se enuncia das seguintes formas:

H_0 : os grupos são iguais, isto é, há uma probabilidade igual de surgir um indivíduo com a mesma doença no grupo de controlo e no grupo de doentes

ou

H_0 : os dois grupos não estão associados ou são independentes.

A sua hipótese alternativa estipula que:

H_1 : os grupos são diferentes, o que equivale a dizer que há uma associação entre os dois grupos

ou que

H_1 : os dois grupos são dependentes.

A regra de decisão do teste (*decision rule*) estipula que:

- Se o valor de RC observado toma um valor unitário admite-se como verdadeira a hipótese nula, o que significa que os dois grupos são independentes, ou que não rejeitamos a hipótese nula.
- Se rejeitarmos a hipótese nula e o valor obtido for inferior à unidade (mas superior a zero) diz-se que os dois grupos estão associados e que essa associação é negativa entre eles.
- Se rejeitarmos a mesma hipótese nula e o valor de RC for superior à unidade então diz-se que os dois grupos estão associados e que essa associação é positiva entre eles.

Do que se disse se depreende que os rácio de controlo, O_c , e o “odds ratio”, OR, ou rácio dos produtos cruzados, RC, têm nesta metodologia papéis cruciais.

Igualmente importante é o Intervalo de Confiança (IC) de RC ou Intervalo de Aceitação da hipótese nula, cuja probabilidade é previamente fixada; este intervalo está associado ao intervalo de rejeição pois que a reunião de ambos vem igual ao intervalo $]-\infty +\infty[$; este último intervalo corresponde à probabilidade que se quer correr de tomar uma decisão errada, isto é, de rejeitar uma hipótese que afinal é correcta. Para construir este intervalo fixa-se um nível de significância α , geralmente, 5% ou 1%, sendo esta a probabilidade que se quer correr de rejeitar uma hipótese válida ou verdadeira; dada a interligação entre os dois intervalos esta afirmação equivale a “aceitar” (não rejeitar) H_0 com uma confiança de $100-\alpha = 95\%$ ou 99% .

Nesse intervalo de confiança intervêm para além do “odds ratio”, OR, que toma o papel de valor central do intervalo, um produto de dois factores: o erro padrão da estimativa e o valor que a tabela da t de student ou da normal reduzida – consoante o caso – dá para esse risco ou confiança. Esse produto vai depois deduzir-se e adicionar-se ao “odds ratio”, OR, definindo assim a região de aceitação – assim chamada por permitir ‘aceitar’ ou preferivelmente ‘não rejeitar’ a hipótese nula H_0 – do teste. A região de rejeição, que tem probabilidade dada pelo nível de significância α , fica assim quase automaticamente delimitada por esses mesmos pontos extremos. O que se disse pode traduzir-se em linguagem estatística quer por (1), se definido em termos de intervalo de rejeição,

$$P[\lambda < \theta - Z_{\alpha/2}^* \sigma_\theta, \quad \lambda > \theta + Z_{\alpha/2}^* \sigma_\theta] = \alpha \quad (1)$$

e (1’), se definido em termos de intervalo de aceitação,

$$P[\theta - Z_{\alpha/2}^* \sigma_\theta \leq \lambda \leq \theta + Z_{\alpha/2}^* \sigma_\theta] = 1 - \alpha \quad (1')$$

onde θ é o valor da taxa de risco, OR, observada na amostra, σ_θ é o erro padrão dessa estimativa ou taxa, $Z_{\alpha/2}$ e $Z_{1-\alpha/2}$ são os valores que se lêem da tabela uma vez fixado o nível de significância α ou de confiança $1-\alpha$, e λ é a taxa de risco (desconhecida) a estudar.

Como nem sempre a estatística a utilizar para estimar o rácio de risco, RR, é simétrica – o que se explica, por exemplo, porque, as taxas nunca são negativas e o seu valor se distribui no intervalo $[0 a +\infty[$, e ainda, como se pode ver facilmente, o intervalo $[0 1[$ concentra os valores de RR correspondentes a associações negativas –, torna-se necessário, nestes casos, introduzir algumas transformações sobre os dados iniciais de molde a controlar esses factores de assimetria (e também de não normalidade). A transformação a introduzir no contexto que estamos a desenvolver é a operação de logaritmização, operação que, por razões que nos dispensamos de apresentar neste momento, ocupa, aqui, um lugar de destaque.

O que se disse faz com que as duas probabilidades tenham que ser escritas ou como (2), se em termos de região de aceitação do logaritmo do rácio de risco,

$$P[\ln \theta - z_{\alpha/2} \sigma_{\ln \theta} \leq \ln \lambda \leq n\theta + z_{1-\alpha/2} \sigma_{\ln \theta}] = 1 - \alpha \quad (2)$$

ou como (2'), se em termos de região de rejeição do mesmo logaritmo

$$P[\ln \lambda > \ln \theta - z_{\alpha/2} \cdot \sigma_{\ln \theta}, \quad \ln \lambda > \ln \theta + z_{1-\alpha/2} \cdot \sigma_{\ln \theta},] = \alpha \quad (2')$$

o que é equivalente, em termos de rácio de risco, λ , a ter (3), se em termos de região de rejeição,

$$P\left[\lambda < \theta e^{-z_{\alpha/2}^* \sigma_{\ln \theta}}, \quad \lambda > \theta e^{z_{1-\alpha/2}^* \sigma_{\ln \theta}}\right] = \alpha \quad (3)$$

ou (3'), se em termos de região de aceitação,

$$P\left[\theta e^{-z_{\alpha/2} \sigma_{\ln \theta}} \leq \lambda \leq \theta e^{z_{1-\alpha/2}^* \sigma_{\ln \theta}}\right] = 1 - \alpha \quad (3')$$

A estimação do erro padrão da estimativa depende do facto de a taxa de risco λ referida ser uma razão de risco, uma razão de proporções ou um caso-controlo. No primeiro caso o erro padrão obtém-se recorrendo à expressão (4)

$$\sigma_{\ln \theta} = \sqrt{\frac{1}{a} + \frac{1}{b}} \quad (4)$$

onde a e b são, respectivamente, o número de doentes com o factor e o número de doentes sem o factor. No segundo caso o erro padrão obtém-se recorrendo à expressão (5)

$$\sigma_{\ln \theta} = \sqrt{\frac{c}{aN_1} + \frac{d}{bN_0}} \quad (5)$$

onde a , b , c , d , N_0 e N_1 são, respectivamente, o número de doentes com o factor, o número de doentes sem o factor, o número de indivíduos do grupo de controlo com o factor, o número de indivíduos do mesmo grupo sem o factor, o número total de elementos do grupo de controlo ($N_0=c+d$), e o número total de doentes ($N_1=a+b$).

No terceiro e último caso, conhecido como casos-controlo, pode estimar-se recorrendo à expressão (6)

$$\sigma_{\ln t} = \sqrt{1/a + 1/b + 1/c + 1/d} \quad (6)$$

Em termos práticos o teste processa-se da seguinte forma: uma vez calculado o intervalo de confiança se esse intervalo contiver o valor unitário então admite-se como verdadeira a hipótese nula, isto é, que as duas amostras (a de controlo e a dos doentes) são iguais e independentes, ou que ambos os grupos têm as mesmas características. Se não contém o valor unitário então rejeita-se H_0 e diz-se que os dois grupos são dependentes e por isso que há associação; neste caso temos duas alternativas: quando o valor do "odds ratio", OR, é inferior à unidade e o intervalo de confiança do rácio de risco não contém o valor unitário, caso em que se diz que a *associação entre os dois grupos é do tipo negativo*, e quando o valor de OR é superior à unidade e o intervalo de confiança não engloba o valor um, caso em que se diz que se está em presença de uma *associação positiva entre ambos os grupos*.

3. POPULAÇÃO E AMOSTRA

O universo ou população é constituído pelos jovens (rapazes e raparigas) dos 12 aos 16 anos oriundos de um distrito do interior Norte (Vila Real).

A amostra integra 907 raparigas e 905 rapazes, num total de 1812 jovens, que foram inquiridos em 6 escolas localizadas em três tipos de zonas: urbana, suburbana e rural. As escolas e os jovens foram seleccionados por processos que garantem a aleatoriedade entre as escolas e os membros seleccionados para que se possa extrapolar as conclusões a todo o universo ou população.

4. RESULTADOS OBTIDOS

O estudo inicia-se através da abordagem dos factores determinantes do consumo de tabaco para todos os jovens dos 12 aos 16 anos, isso é, para os jovens dos sexos masculino e feminino; leva depois a cabo o mesmo tipo de análise para cada um dos sexos individualmente, concluindo-se pela análise das diferenças entre rapazes e raparigas.

Mas antes disso, porém, façamos uma breve análise descritiva da amostra.

4.1 CARACTERIZAÇÃO DA AMOSTRA

4.1.1 Prevalência e Precocidade Tabágicas

O número total de indivíduos da amostra é de 1812 adolescentes de ambos os sexos com idades compreendidas entre os 12 e os 16 anos. A origem geográfica dos jovens inquiridos é o distrito de Vila Real (Portugal), repartidos pelo interior das zonas urbanas (origem urbana), das zonas suburbanas e das zonas rurais.

Do total referido há 63.4% que declararam nunca terem fumado e 36.6% que declararam já terem fumado pelo menos uma vez. Destes há 67% que afirmaram ter apenas experimentado; os restantes 33% declararam ter fumado nos últimos 30 dias. Assim a prevalência de fumadores – definida esta como a percentagem de jovens que fumaram nos últimos 30 dias – é de 12.3% do total.

Dos jovens que alguma vez experimentaram fumar a grande maioria (98.5%) experimentou antes dos 15 anos de idade, e destes cerca de 2/3 (66.3%) experimentaram mesmo antes de completarem 13 anos.

A amostra prova que os jovens do sexo masculino se iniciam no consumo de tabaco por volta dos 12.5 anos, cerca de meio ano mais cedo do que as jovens do sexo feminino. Com uma certeza de 95% pode afirmar-se que os adolescentes rapazes se iniciaram entre os 10.8 e os 14.2 anos (com uma amplitude do intervalo de confiança de 3.4 anos), e que as adolescentes raparigas o fizeram entre os 11.7 e os 14.3 anos (com uma amplitude do intervalo de confiança de 2.6 anos). Do que se deixou entre parêntesis se conclui que as idades de iniciação são mais concentradas nas raparigas do que nos rapazes, dando a ideia que apesar de começarem cerca de meio ano mais tarde não pretendem atrasar-se mas antes recuperar o *tempo perdido*.

Um facto curioso que se retira desta amostra é que os adolescentes que reprovam se iniciam mais tarde no consumo de tabaco (13.0 para os Homens, H, e 13.3 para as Mulheres, M) do que entre os que nunca reprovaram (12.2 H e 12.9 M, respectivamente), o que significa em qualquer dos casos que os melhores estudantes – por não reprovarem – são igualmente mais precoces na iniciação ao tabaco.

Pode afirmar-se que a prática de alguma religião atrasa a data de iniciação ao tabaco, em média, em 2.4 meses para os rapazes e em 1.2 meses para as raparigas.

O número de horas de sono não parece ter grande influência na idade de iniciação ao fumo: apenas nas jovens se verifica um atraso de 1.2 meses para as que dormem mais (de 8 horas), o que de alguma forma permite afirmar que quem dorme mais chega *mais tarde* ao vício do tabaco.

As idades de iniciação ao tabaco variam consoante o local de residência dos adolescentes; os mais precoces são os que vivem em zonas suburbanas (12.1 para os H e 12.9 para as M) e os menos precoces são os das zonas rurais (12.7 para os H e 13.1 para as M).

O factor escolaridade dos pais parece não ser factor importante na idade de iniciação ao tabaco. Pelo menos para os rapazes (diferença máxima média de 1.2 meses). No caso das raparigas há maiores diferenças: 0.4 anos (ou 4.8 meses) em média. Pode ainda concluir-se, a este propósito, que a idade de iniciação ao tabaco

umenta à medida que se reduz o número de anos de escolaridade dos pais, isto no caso das raparigas.

Quando os pais delas vivem separados as jovens aceleram ligeiramente (1.2 meses em média) a idade de iniciação no tabaco. No caso dos rapazes não se regista qualquer alteração digna de registo.

O facto de ambos os pais ou um só consumir tabaco é factor que acelera a sua própria idade de iniciação, isto em ambos os sexos. Se apenas um dos pais consome a idade média de iniciação ao tabaco sofre um atraso de meio ano (ou 6 meses) tanto no caso dos rapazes como no caso das raparigas.

A existência de amigos fumadores acelera ligeiramente (2.4 meses) a idade de iniciação ao tabaco dos rapazes e não parece ter efeito no caso das raparigas.

Finalmente, quando os jovens têm depressões de origem psíquica, no caso dos rapazes têm tendência a refugiarem-se mais cedo no tabaco; no caso das raparigas este factor não introduz qualquer alteração.

4.1.2 Factores determinantes da precocidade do consumo de tabaco dos jovens

De acordo com a amostra recolhida de todos os factores referidos os que, por ordem decrescente de importância, mais antecipam a idade média de iniciação dos rapazes ao consumo de tabaco são: em 1º lugar o residir em ambiente suburbano, em 2º, *ex-aequo*, o ambos os pais fumarem, e, curiosamente, o ter aproveitamento escolar, em 3º, *ex-aequo*, o não terem religião e sofrerem de depressão, e em 4º, igualmente *ex-aequo*, a escolaridade dos pais ser superior a 9 anos e um dos pais fumar. No caso das raparigas os factores que aceleram a idade da iniciação tabágica são: em 1º lugar ambos os pais fumarem, em 2º a escolaridade dos pais ser superior a 9 anos, e em 3º *ex-aequo*, não ter reprovado, não ter religião, viver em ambiente suburbano e os pais estarem divorciados.

4.1.3 Opinião dos jovens acerca de diversos assuntos ligados com o consumo de tabaco

Auscultados sobre diversas questões ligadas com o consumo de tabaco os adolescentes emitiram as opiniões que se deixam em seguida:

- Praticamente a totalidade (97.5%) dos fumadores têm a opinião que fumar é prejudicial para a saúde; contudo, esta opinião têm-na 'apenas' 88.8% dos não fumadores.
- 62.6% dos não fumadores afirmam que o consumo do tabaco os faz sentir mal; curiosamente apenas 41.3% dos fumadores afirmam o mesmo.

- Há ainda uma percentagem de 13% de fumadores que afirmam que o hábito de fumar é necessário para viver melhor, e apenas 7.1% têm a mesma opinião entre os não fumadores.
- Dos fumadores há 66% que afirmam ser o tabaco uma dependência difícil de superar; dos não fumadores 58.7% concordam com esta mesma opinião.
- Apesar de tudo uma percentagem bastante elevada (60.5%) dos adolescentes fumadores considera que o tabaco é uma forma errada de reagir a situações que causam nervosismo; a percentagem dos não fumadores que têm a mesma opinião é 75.1%.

4.1.4 Factores determinantes do consumo de tabaco

Os dados mostram que a prevalência do consumo de tabaco é inferior para as jovens raparigas (9.2%) do que para os rapazes (15.5%).

Independentemente do sexo a percentagem de fumadores aumenta com a idade (referimo-nos apenas à faixa etária dos 12 - 16 anos), passando a prevalência de 2.0%, aos 12 anos, para 28.4%, aos 16 anos; a prevalência média global é 12.3% como já se referiu anteriormente.

A prevalência aumenta com o absentismo às aulas: passa de 8.1% quando não há absentismo nenhum para 39.1% quando há absentismo.

Entre os inquiridos a percentagem de adolescentes que já tinham reprovado pelo menos uma vez era de 30.6%. Os fumadores com sucesso escolar eram 10% mas com insucesso escolar eram 17.5%.

O consumo de tabaco aumenta com o número de filhos do casal (fratria) passando de 9.9% no caso de uma família com apenas um filho para 15.1% no caso de uma família com três ou mais filhos.

O consumo de tabaco é igualmente influenciado pela forma de vida (em comunhão ou não) dos pais, passando de 11.8% quando os pais vivem juntos para 17.1% quando vivem separados.

Esta amostra prova que os adolescentes que praticam uma religião (81.3%) em geral são menos fumadores (10.7% para os que se dizem praticantes de uma religião contra 19.3% para os que se declaram não praticantes ou ateus).

A amostra revela que o número de horas de sono está associado ao consumo de tabaco: os que dormem menos de 8 horas/dia apresentam uma proporção mais elevada de fumadores (18.7% contra 9.8%).

O vício do tabaco aumenta quando se passa de adolescentes cujos pais não fumam para outros em que um dos pais fuma e para outros ainda em que ambos fumam. As percentagens são, respectivamente, 10.3%, 13.4% e 20.3%.

A influência dos amigos é importante no que diz respeito ao consumo de tabaco; o facto de ter amigos fumadores passa a percentagem de fumadores de menos de 1% para 15%.

Igualmente importante para explicar o consumo de tabaco é o facto de ter namorado/a, e, neste caso, de ele/a ser fumador/a; se não veja-se: se o/a jovem não tem namorado/a a percentagem de fumadores é 8%, se tem namorado/a passa a 14.4%, mas se este/a for fumador/a então a percentagem passa a 34.3%.

O consumo de álcool é igualmente um factor importante a ter em atenção na explicação do consumo de tabaco; de facto, há apenas 4.8% de adolescentes que não consomem álcool e são fumadores mas há 32.6% que consomem álcool e fumam.

A existência de depressão mais do que duplica a prevalência de fumadores: se o adolescente não sofrer de depressão há 10% de probabilidades de fumar, mas se sofrer de depressão essa percentagem passa a 20.8%.

Há duas situações em que não parece haver regras bem definidas na explicação do consumo de tabaco: uma tem que ver com o local de residência e outra com a escolaridade dos pais.

Qual é a importância do local de habitação na explicação da prevalência do consumo de tabaco? Os dados mostram que esta cresce à medida que passamos do meio suburbano (10.8%), para o urbano (12.1%) e deste para o rural (13.2%). Como explicar este último valor e o facto de ele ser maior que os anteriores? Será que é uma questão de maior falta de informação por parte dos rurais?

Vejamos agora qual a importância da escolaridade dos pais? Os dados mostram que a probabilidade de ser fumador um indivíduo cujos pais têm entre 5 e 9 anos de escolaridade é menor, 11.4%, mas que ela passa para 12.5% se os pais tiverem mais de 9 anos de escolaridade e para 12.6% se tiverem até à 4ª classe. Em função disso pode afirmar-se que não parece existir qualquer nexo entre os dois atributos.

4.2 INFERÊNCIA (BIO)ESTATÍSTICA

De referir, antes de iniciar este novo tipo de análise, que no que se vai seguir se considerou sempre um nível de significância de 5%, o que significa que é este o risco que voluntariamente queremos correr de rejeitar uma hipótese válida.

Como já se referiu anteriormente vamos fazer esse tipo de inferência primeiro para a amostra no seu todo, depois para o sexo feminino, em seguida para o sexo masculino e finalmente apreciaremos as diferenças entre os dois sexos.

4.2.1 Ao nível global (ambos os sexos tomados conjuntamente)

1. *Efeito da escola no consumo de tabaco*: em geral há uma associação negativa entre o vício de fumar e as escolas que o aluno(a) frequenta uma vez que os rácios OR se situam entre 0 e 1, e, além disso, 4 de 5 intervalos de confiança de 95% situam-se integralmente entre 0 e 1 também. Apenas

na escola nº 2 nos ficam algumas dúvidas. A escola nº 1 foi considerada a escola de controlo.

2. *Efeito da idade*: há uma relação positiva entre o hábito de fumar dos jovens e a idade, isto é, à medida que a idade aumenta o vício do tabaco aumenta também. O grupo de controlo considerado foi o dos 12 anos.
3. *Efeito do absentismo*: também o absentismo dos jovens às aulas é factor que está associado positivamente ao consumo de tabaco, isto é, os jovens consomem tanto mais tabaco quanto maior for o seu absentismo às aulas. O grupo de controlo é constituído pelos indivíduos que não faltam às aulas.
4. *Efeito do mau aproveitamento escolar*: há uma associação positiva entre insucesso escolar e hábito de fumar dos jovens. O grupo de controlo integra os indivíduos que não têm problemas de insucesso escolar.
5. *Efeito do local de residência*: não se pode garantir uma associação entre o hábito de fumar e o local de residência dos jovens. O grupo de controle integra os jovens com residência urbana.
6. *Efeito do número de filhos por agregado familiar*: também não se pode afirmar que haja associação entre o hábito de fumar e o número de filhos por agregado familiar. O grupo de controle é constituído pelos jovens oriundos de famílias com apenas um filho.
7. *Efeito da escolaridade dos pais*: não há associação entre o hábito de fumar dos jovens e a escolaridade dos pais. O grupo de controle é constituído pelos jovens cujos pais têm apenas de 0 a 4 anos de escolarização.
8. *Efeito do estado conjugal ou civil dos pais*: há uma associação positiva entre hábito de fumar dos jovens e o facto de os pais estarem separados. O grupo de controle utiliza como referência os jovens cujos pais vivem juntos.
9. *Efeito da religião*: há uma associação negativa entre o atributo fumar e o efeito religião, isto é, à medida que passamos de jovens que não seguem qualquer religião para os que seguem uma tende a decrescer a gravidade do fenómeno. O grupo de referência é o daqueles que não praticam qualquer religião.
10. *Efeito do número de horas de sono*: é também negativa a relação do fenómeno com o número de horas de sono. O grupo de controlo é o daqueles que dormem menos de 8 horas.
11. *Efeito do consumo de tabaco por parte dos pais*: a associação é positiva com o facto de ambos os pais serem fumadores e duvidosa se apenas um deles fumar. O grupo de referência é o daqueles cujos pais não fumam.
12. *Efeito do consumo de tabaco por parte dos amigos*: há também uma associação bastante forte e positiva entre hábito de fumar dos jovens e o facto de terem amigos fumadores. O grupo de comparação é o daqueles que não têm amigos fumadores.
13. *Efeito do consumo de tabaco por parte do namorado*: há igualmente uma forte relação positiva entre o consumo de tabaco por parte dos jovens e o

facto de o namorado ser fumador; esta conclusão sai reforçada se se disser que essa relação é negativa se o jovem não tiver namorado. O grupo de comparação é o daqueles cujo namorado não fuma.

14. *Efeito do consumo de bebidas*: a relação é fortemente positiva. O grupo de controlo é o daqueles que não consomem álcool.

15. *Efeito de estados depressivos*: é igualmente positiva a relação entre o vício do tabaco e a depressão dos jovens. O grupo de comparação é o daqueles que não são depressivos.

4.2.2 Individuais femininos

1. *Efeito da escola no consumo de tabaco*: o rácio dos produtos cruzados, ou seja o estimador do rácio de risco, situa-se, em média, entre 0.3 e 0.5; os intervalos de confiança obtidos permitem-nos assegurar, considerando os jovens da escola 1 como grupo de controlo – o que, apesar de o fazermos, nos merece algumas reservas –, que há efectivamente associação entre escola e consumo de tabaco, com excepção da nº 6, e que essa associação é do tipo negativo. Na escola nº 6 rejeitou-se a hipótese nula pelo que se pode dizer que há independência entre estes dois elementos ou atributos.
2. *Efeito da idade*: rejeita-se a hipótese nula para as idades 14 a 16 o que significa que há uma associação positiva entre o vício de fumar e a idade; não se confirma essa associação para a idade de 13 anos. As jovens de 12 anos foram consideradas como grupo de controlo.
3. *Efeito do absentismo às aulas*: os valores encontrados indicam uma relação fortemente positiva entre o vício do fumo e o absentismo às aulas por parte das raparigas.
4. *Efeito do mau aproveitamento escolar*: os valores encontrados não nos permitem afirmar que há qualquer associação, para as raparigas, entre fumar e insucesso escolar.
5. *Efeito do local de residência*: os dados também não confirmam alguma associação entre local de residência e hábito de fumar por parte das jovens dos 12 aos 16 anos.
6. *Efeito do número de filhos por agregado familiar*: também o número de filhos do agregado familiar não parece ter qualquer associação com o hábito de fumar do sexo feminino.
7. *Efeito da escolaridade dos pais*: há igualmente independência entre este vício e a escolaridade dos pais.

8. *Efeito do estado conjugal ou civil dos pais*: o estado conjugal dos pais também não tem associação com o vício do tabaco das jovens.
9. *Efeito da religião*: já se conclui haver uma associação do tipo negativo entre hábitos religiosos e de fumo, pelo que jovens que praticam alguma religião têm tendência a fumar menos.
10. *Efeito do número de horas de sono*: como o extremo superior do intervalo de confiança é praticamente 1 podemos dizer que há uma associação negativa a tender para nula entre aquele hábito e o número de horas de sono (+ de 8h); pode assim dizer-se que as jovens que dormem menos são as que mais fumam.
11. *Efeito do consumo de tabaco por parte dos pais*: verifica-se uma associação positiva (a tender para nula ou independente) entre hábitos tabágicos das jovens e o facto de ambos os pais serem fumadores, pelo que se pode concluir que se ambos os pais fumam as filhas tendem a fumar mais. Quando um deles fuma já não se confirma essa relação talvez pelo efeito que o outro possa exercer sobre a jovem.
12. *Efeito do consumo de tabaco por parte dos amigos*: há uma forte associação positiva entre o fumar das jovens e o fumar dos seus amigos.
13. *Efeito do consumo de tabaco por parte do namorado*: há igualmente uma forte associação positiva entre o vício do fumo das jovens e o vício do fumo por parte do namorado. Se a jovem não tem namorado não tem tanta tendência a fumar (neste caso há independência entre ambos os factores).
14. *Efeito do consumo de bebidas alcoólicas*: também o consumo de bebidas alcoólicas por parte das jovens está fortemente associado, positivamente, com o consumo de bebidas alcoólicas. Quanto mais bebem as jovens mais elas fumam.
15. *Efeito de estados depressivos*: o vício feminino do tabaco está igualmente positivamente correlacionado com o estado depressivo das jovens.

4.2.3 Individuais masculinos

1. *Efeito da escola no consumo de tabaco*: embora haja 2 escolas (as números 2 e 3) em que parece não haver associação nas outras parece haver uma associação do tipo negativo.
2. *Efeito da idade*: há uma associação positiva entre os factores idade (a partir dos 13 anos) e hábito de fumar dos rapazes.
3. *Efeito do absentismo*: há uma relação forte (positiva) entre o absentismo e o hábito de fumar dos rapazes.
4. *Efeito do mau aproveitamento escolar*: também é positiva a sua relação com o mau aproveitamento escolar.

5. *Efeito do local de residência*: não há associação ou há independência entre estes dois factores para os jovens do sexo masculino.
6. *Efeito do número de filhos do agregado familiar*: há independência com o número de filhos do agregado familiar.
7. *Efeito da escolaridade dos pais*: há igualmente independência com a escolaridade dos pais.
8. *Efeito do estado conjugal ou civil dos pais*: se os pais estão separados os filhos tendem a fumar mais, o que quer dizer que há associação e que ela é positiva.
9. *Efeito da religião*: confirma-se também nos homens a tendência para uma associação do tipo negativo (mas próxima da independência) entre fumar dos jovens e a prática de uma religião.
10. *Efeito do número de horas de sono*: os valores confirmam que o jovem que dorme mais de oito horas por dia fuma menos (associação negativa).
11. *Efeito do consumo de tabaco por parte dos pais*: confirma-se haver correlação positiva entre fumo dos filhos e fumo dos pais.
12. *Efeito do consumo de tabaco por parte dos amigos*: o mesmo se passa neste caso. A relação é altamente positiva.
13. *Efeito do consumo de tabaco por parte do namorado*: é também forte a correlação entre estes dois factores. Mas se a namorada não fumar o jovem tem tendência a não fumar ou a fumar menos.
14. *Efeito do consumo de bebidas alcoólicas*: há uma relação fortemente positiva com o consumo de bebidas alcoólicas.
15. *Efeito de estados depressivos*: depressão e consumo de tabaco estão correlacionados positivamente.

4.2.4 Diferenças entre os dois sexos

1. *Efeito da escola no consumo de tabaco*: em geral há concordância entre ambos os sexos: a correlação escola – hábitos tabágicos é negativa com uma escola ou outra em que não se prova a associação.
2. *Efeito da idade*: a associação é positiva em qualquer dos casos.
3. *Efeito do absentismo*: a associação é positiva; há concordância entre os dois sexos.
4. *Efeito do mau aproveitamento escolar*: detecta-se alguma discrepância entre os dois sexos: associação positiva no caso dos rapazes e nula no caso das raparigas.

5. *Efeito do local de residência*: não há associação entre os dois factores, havendo concordância entre os dois sexos.
6. *Efeito do número de filhos por agregado familiar*: confirma-se que há independência entre estes factores tanto para rapazes como para raparigas.
7. *Efeito da escolaridade dos pais*: há independência entre estes factores tanto para rapazes como para raparigas.
8. *Efeito do estado conjugal ou civil dos pais*: não há concordância em ambos os sexos: a associação é positiva no caso dos rapazes mas nula (independência) no caso das raparigas.
9. *Efeito da religião*: há associação negativa: neste caso chega-se praticamente à mesma conclusão tanto com rapazes (quase independência) como com raparigas.
10. *Efeito do número de horas de sono*: associação negativa (a quase nula no caso das raparigas). Pode dizer-se que há concordância entre os dois sexos.
11. *Efeito do consumo de tabaco por parte dos pais*: prova-se haver associação positiva para os rapazes e independência para as raparigas.
12. *Efeito do consumo de tabaco por parte dos amigos*: a associação é positiva para os rapazes, as raparigas e o conjunto de ambos os sexos.
13. *Efeito do consumo de tabaco por parte do namorado*: sendo o namorado/a fumador/a o/a jovem fuma mais seja ele rapaz ou rapariga; já quando não tem namorado/a há independência para as raparigas e associação negativa para os rapazes.
14. *Efeito do consumo de bebidas alcoólicas*: há associação positiva e há concordância entre ambos os sexos.
15. *Efeito de estados depressivos*: é positiva a relação; há igualmente concordância entre os jovens rapazes e raparigas.

6. CONCLUSÕES

Do que se deixa escrito se conclui que o problema do consumo de tabaco continua a ser um problema com alguma gravidade, tanto mais que os jovens continuam a consumir tabaco em número relativamente elevado; que a idade de iniciação a este consumo é em geral muito baixa, a maioria das vezes antes dos 13 anos; que esse problema é grave nas raparigas e nos rapazes; que os rapazes em geral são mais precoces mas que as raparigas não sendo tanto aceleram depois ficando praticamente ao mesmo nível; que o facto de reprovar parece estar associado a um atraso de na idade de iniciação ao tabaco; que a prática de uma religião dificulta a penetração do vício de fumar; que o dormir pouco acelera o consumo; que os jovens das zonas sub-urbanas são os mais precoces e os das zonas rurais os que mais tarde chegam ao contacto com o tabaco; que o número de anos de escolaridade dos pais parece não ter influência sobre o início do consumo; que o facto de os pais viverem separados é factor de aceleração ou de precocidade; que o facto de um ou ambos os pais fumarem acelera a idade de consumo de tabaco; que a existência de amigos e

sobretudo de namorados fumadores são factores que antecipam igualmente a idade de consumo de tabaco; que os jovens com depressões psíquicas caem mais rapidamente nas malhas do tabaco.

Aplicando directamente a metodologia defendida pela OMS e transcrita em Ramalheira e Cardoso (1995) conclui-se que naquele distrito do norte do país há uma associação positiva entre hábitos tabágicos dos jovens no seu todo e os seguintes factores: idade (dos 12 aos 16 anos), absentismo às aulas, o insucesso escolar, os pais estarem separados, hábitos tabágicos dos pais, amigos e namorados, o consumo de bebidas alcoólicas e estados depressivos dos jovens. Por sua vez há associação negativa entre consumo de tabaco dos jovens e a escola de frequência (?), a prática de uma religião e o número de horas de sono. Por fim, não pode assegurar-se que existe associação entre o fumo dos jovens e os factores: o local de residência dos jovens, o número de filhos do agregado familiar e a taxa de escolarização dos pais.

Para terminar esta conclusão podemos garantir que há *efeitos coincidentes (do tipo positivo)* entre o hábito de fumar de ambos os sexos e a idade, o absentismo às aulas, o consumo de tabaco por parte dos amigos, o facto de o/a namorado/a ser fumador/a, o consumo de bebidas alcoólicas, e o facto de o/a jovem sofrer de estados depressivos. Há *efeitos coincidentes (negativos)* entre os dois sexos no que diz respeito à escola de frequência, à prática de uma religião, ao número de horas de sono. Há *efeitos coincidentes (não associação ou independência)* entre ambos os sexos no que diz respeito ao local de residência, ao número de filhos do agregado familiar, e à taxa de escolarização dos pais. Há *efeitos conflituantes ou não coincidentes* entre rapazes e raparigas no que diz respeito aos factores mau aproveitamento escolar (associação positiva para os rapazes (H) e independência para as raparigas (M)), à vivência em comum dos pais (positiva para os jovens do sexo masculino e independência para as jovens do sexo feminino), e ao facto de não ter namorado/a (negativa para os rapazes e nula ou independência para as raparigas).

7. BIBLIOGRAFIA

- Quadros *et alii* – “Exercícios de Bioestatística”, Laboratório de Cálculo, Universidad de Barcelona, 19
- Ramalheira e Cardoso – “”, ed. , 1995
- Manso, J. R: Pires – “Curso de Estatística”, 2^a ed, ed. Autor, 1993
- Murteira, B. Ferreira – “Teoria das Probabilidades e Estatística Matemática”, 2^a ed, MacGraw-Hill, 199
- Siegel, Sidney – “Estatística no paramétrica aplicada a las ciências de la conducta”, Trillas, Biblioteca de Psicologia, Cneip, 1976

Silva, C. M. da – “Estatística Aplicada à Psicologia e às Ciências Sociais”, McGraw-Hill, 1994

Spiegel, Murray R. – “Estatística”, 3ª ed., MacGraw-Hill, Makron Books, 1993

Vicente, Paula *et alii* – “Sondagens – a Amostragem como Factor Decisivo de Qualidade”, Edições Sílabo, 1996

ANEXOS

A.1. CONSUMO DE TABACO - SEXO FEMININO

CONSUMO DE TABACO - S FEMININO					
ESCOLA	S FEMININO		RC	LIM.INF	LIMSUP
	NAO FUMA	FUMA	OR		
1	164	31	1*		
2	98	7	0,37788	0,160301	0,890783
3	124	8	0,341311	0,151608	0,768386
4	180	16	0,470251	0,248117	0,891257
5	187	14	0,398067	0,203677	0,770186
6	71	7	0,521581	0,219355	1,240211
IDADE					
12	93	1	1*		
13	247	12	4,518219	0,579376	35,23498
14	256	32	11,625	1,566231	86,28397
15	168	21	11,625	1,538954	87,81332
16	60	17	26,35	3,416778	203,2097
ABSENTISMO AULAS					
NAO	747	40	1*		
SIM	76	43	10,56612	6,466688	17,2643
INSUCESSO ESCOLAR					
NAO	600	61	1*		
SIM	224	22	0,966042	0,579509	1,610393
RESIDENCIA					
URBANO	386	43	1*		
SUB-URBANO	101	10	0,888787	0,431694	1,829866
RURAL	335	30	0,803888	0,493131	1,310474
FRATRIA					
UNICO	53	6	1*		
1º FILHO	284	21	0,653169	0,251718	1,694869
2º FILHO	283	28	0,873969	0,345083	2,213445
3º E +	179	24	1,184358	0,460031	3,049151
ESCOLARIDADE PAIS (ANOS)					
0-4	401	39	1*		
5-9	135	15	1,14245	0,610523	2,141708
+ DE 9	285	29	1,046244	0,632036	1,718201
EST CIVIL PAIS					
JUNTOS	735	72	1*		
SEPARADOS	89	11	1,261704	0,644623	2,4695
PRÁTICA DE RELIGIAO					
NAO	125	28	1*		
SIM	694	55	0,353798	0,216038	0,567175
HORAS DORMIDAS					
- DE 8	182	26	1*		
8 OU MAIS	641	56	0,611544	0,373372	1,001645
PAIS FUMADORES					
NAO	432	38	1*		
UM	326	34	1,185664	0,73036	1,924801
OS DOIS	63	11	1,984962	0,964978	4,083073
AMIGOS FUMADORES					
NAO	147	2	1*		
SIM	672	81	8,859375	2,153804	36,44182
NAMORADO FUMADOR					
NAO	183	13	1*		
SIM	91	27	4,176669	2,347745	8,476736
NAO TEM NAMOR.	545	43	1,110656	0,677907	2,111746
CONSUMO ALCOOL					

NÃO	691	29	1*		
SIM	125	53	10,1029	6,182087	16,51037
DEPRESSÃO					
NÃO	647	46	1*		
SIM	177	37	2,940187	1,849309	4,725578

A2: CONSUMO DE TABACO - S MASCULINO

ESCOLA	CONSUMO DE TABACO - S MASCULINO				
	SEXO MASCULINO			LIM.INF	LIMSUP
	NÃO FUMA	FUMA	OR		
1	185	47	1*		
2	82	13	0,624027	0,320286791	1,215816
3	100	30	1,180851	0,703006856	1,993493
4	151	22	0,573482	0,330851897	0,994044
5	150	21	0,551064	0,315492856	0,96253
6	97	7	0,284054	0,123711878	0,652212
IDADE					
12	101	3	1*		
13	217	20	3,102919	0,901295346	10,68252
14	252	37	4,943122	1,490304711	16,39561
15	139	51	12,35252	3,749339949	40,69642
16	56	29	17,43452	5,082058567	59,81092
ABSENTISMO AULAS					
NÃO	689	86	1*		
SIM	75	54	5,768372	3,807334622	8,739478
INSUCESSO ESCOLAR					
NÃO	531	65	1*		
SIM	234	75	2,618343	1,81578216	3,77563
RESIDÊNCIA					
URBANO	358	59	1*		
SUB-URBANO	122	17	0,845513	0,474667781	1,506088
RURAL	279	63	1,370148	0,92947051	2,019757
FRÁTRIA					
ÚNICO	76	11	1*		
1º FILHO	280	41	1,011686	0,49630723	2,062257
2º FILHO	261	54	1,429467	0,712069772	2,869629
3º E +	120	29	1,669697	0,787781685	3,538909
ESCOLARIDADE PAIS (ANOS)					
0-4	327	66	1*		
5-9	169	24	0,703804	0,425587793	1,167164
+ DE 9	268	50	0,924355	0,618748367	1,366958
EST CIVIL PAIS					
JUNTOS	698	119	1*		
SEPARADOS	66	21	1,86631	1,100571239	3,164823
PRÁTICA DE RELIGIÃO					
NÃO	147	37	1*		
SIM	611	102	0,663246	0,437081652	1,006437
HORAS DORMIDAS					
- DE 8	232	69	1*		
8 OU MAIS	532	71	0,448731	0,311363511	0,646701
PAIS FUMADORES					
NÃO	438	62	1*		
UM	267	58	1,534614	1,040015626	2,264428
OS DOIS	59	20	2,394751	1,350713441	4,245781

AMIGOS FUMADORES					
NÃO	192	1	1*		
SIM	571	139	46,73905	6,493216516	336,4341
NAMORADO FUMADOR					
NÃO	288	66	1*		
SIM	19	30	7,272727	3,824641158	13,82942
NÃO TEM NAMOR.	457	44	0,420131	0,279106304	0,632412
CONSUMO ÁLCOOL					
NÃO	548	33	1*		
SIM	203	106	8,671145	5,682572987	13,23146
DEPRESSÃO					
NÃO	660	103	1*		
SIM	104	37	2,279686	1,484713518	3,500318

A3: CONSUMO DE TABACO: AMBOS OS SEXOS

CONSUMO DE TABACO: AMBOS OS SEXOS					
ESCOLA	NÃO FUMA	FUMA	OR	LIM.INF	LIMSUP
1	349	78	1*		
2	180	20	0,497151	0,294638	0,338856
3	224	36	0,759043	0,497483	1,158122
4	331	36	0,513673	0,338899	0,778579
5	337	35	0,464696	0,303521	0,711457
6	168	14	0,372863	0,205055	0,677999
IDADE					
12	194	4	1*		
13	464	32	3,344828	1,167129	9,585905
14	508	69	6,587598	2,37192	18,29591
15	307	72	11,37459	4,09025	31,63165
16	116	46	19,23276	6,74848	54,81219
ABSENTISMO AULAS					
NÃO	1436	126	1*		
SIM	151	97	7,321139	5,351529	10,01566
INSUCESSO ESCOLAR					
NÃO	1131	126	1*		
SIM	458	97	1,901071	1,428	2,530861
RESIDÊNCIA					
URBANO	744	102	1*		
SUB-URBANO	223	27	0,883144	0,563221	1,384791
RURAL	614	93	1,104809	0,817931	1,492306
FRATRIA					
ÚNICO	129	17	1*		
1º FILHO	564	62	0,834168	0,471897	1,474551
2º FILHO	544	82	1,143815	0,655663	1,995403
3º E +	299	53	1,345072	0,750077	2,412044
ESCOLARIDADE PAIS (ANOS)					
0-4	728	105	1*		
5-9	304	39	0,889474	0,601529	1,317751
+ DE 9	553	79	0,990476	0,724901	1,344608
EST CIVIL PAIS					

JUNTOS	1433	191	1*		
SEPARADOS	155	32	1,548928	1,028538	2,332609
PRÁTICA DE RELIGIÃO					
NÃO	272	65	1*		
SIM	1305	157	0,503436	0,366582	0,691382
HORAS DORMIDAS					
- DE 8	414	95	1*		
8 OU MAIS	1173	127	0,471827	0,353575	0,629626
PAIS FUMADORES					
NÃO	870	100	1*		
UM	593	92	1,349747	0,998147	1,8252
OS DOIS	122	31	2,210656	1,416295	3,45055
AMIGOS FUMADORES					
NÃO	339	3	1*		
SIM	1243	220	20	6,360609	62,88706
NAMORADO FUMADOR					
NÃO	471	79	1*		
SIM	109	57	3,117756	2,0914	4,647799
NÃO TEM NAMOR.	1002	87	0,517661	0,374515	0,71552
CONSUMO ÁLCOOL					
NÃO	1239	62	1*		
SIM	328	159	9,687303	7,050604	13,31004
DEPRESSÃO					
NÃO	1307	149	1*		
SIM	281	74	2,310015	1,699205	3,140391

NADOS VIVOS: ANÁLISE E ESTIMAÇÃO

Autora:
Teresa Bago d' Uva

VOLUME 3

3° QUADRIMESTRE DE 1999

NADOS VIVOS: ANÁLISE E ESTIMAÇÃO

LIVE BIRTHS: ANALYSIS AND ESTIMATION

Autora: Teresa Bago d' Uva

-Gabinete de Estudos e Conjuntura do Instituto Nacional de Estatística

Resumo:

- O objectivo deste estudo é a obtenção de estimativas para o número de nados vivos (de cada um dos sexos) ocorrido por mês em Portugal .

A impossibilidade de quantificar o fenómeno da natalidade oportunamente, devido ao atraso na disponibilidade dos dados, parece não deixar alternativa à previsão.

Serão experimentados alguns métodos de previsão, utilizando modelos dinâmicos univariados (nomeadamente modelos ARIMA e modelos estruturais).

Posteriormente, será estudada a influência de algumas variáveis exógenas através da construção de um modelo econométrico. Espera-se que este modelo econométrico supere os modelos univariados no que diz respeito à qualidade de ajustamento das séries, por ter em conta outras influências para além dos valores passados das variáveis e, como objectivo principal, forneça melhores previsões.

Palavras chave:

- *Natalidade, Modelos ARIMA, Modelos Estruturais, Filtro de Kalman, Modelos Econométricos.*

Abstract:

- The aim of this study is to estimate monthly live births in Portugal (per sex).

The delay in data availability makes us think that they have to be forecasted.

Therefore, there will be tried some forecasting methods making use of univariate models like ARIMA and Structural models.

To study the influence of some exogenous variables, we will build an econometric model. This is expected to provide better results since we take into account information other than just the past values of the dependent variable.

Key-words:

- *Nativity, ARIMA Models, Structural Models, Kalman Filter, Econometric Models.*

1. Introdução

O objectivo deste estudo é a obtenção de estimativas para o número de nados vivos (de cada um dos sexos) ocorrido por mês em Portugal¹.

E porquê a necessidade de estimar valores que são conhecidos com precisão, por ser o registo de todos os nascimentos obrigatório? Essa necessidade surge devido ao atraso que ocorre no apuramento dos dados². Geralmente, é só em Maio de cada ano que se toma como definitivo o nível de nascimentos com vida do ano anterior. No entanto, esse valor é solicitado logo no fim do ano³ quando os dados disponíveis são manifestamente insuficientes (mesmo aqueles meses para os quais já existe alguma informação, estão subavaliados).

A impossibilidade de quantificar o fenómeno da natalidade, oportunamente, parece não deixar alternativa à previsão. As previsões do número de nascimentos são geralmente, feitas em termos de taxas de natalidade (ou de fecundidade) e não em valores absolutos (Shryock e outros, 1976). Porém, os últimos não podem ser deduzidos das primeiras pois isto implicaria o conhecimento da população residente (ou da população feminina em idade fértil) do período em questão.

Serão experimentados alguns métodos de previsão, utilizando modelos dinâmicos univariados (nomeadamente modelos ARIMA e modelos estruturais).

Posteriormente, será estudada a influência de algumas variáveis exógenas através da construção de um modelo econométrico. Espera-se que este modelo econométrico supere os modelos univariados no que diz respeito à qualidade de ajustamento das séries, por ter em conta outras influências para além dos valores passados das variáveis e, como objectivo principal, forneça melhores previsões.

¹ Essas estimativas são necessárias para o cálculo das estimativas mensais da população residente, por sexos, que são obtidas através da equação de concordância:

$$P_t = P_{t-1} + N_t - O_t + SM_t,$$

onde P_t , N_t , O_t e SM_t são, respectivamente, a população, os nados vivos, os óbitos e o saldo migratório do período t (Shryock e outros, 1976).

² Os dados referentes ao número de nados vivos provêm dos verbetes do seu registo. Esse registo deve ser feito no prazo de um mês (é-o, por vezes, mais tarde) na conservatória civil respectiva. Atrasos nas conservatórias levam a que os verbetes demorem algum tempo a chegar ao INE, onde, por sua vez, o seu tratamento e análise não são imediatos, pelo que as estatísticas demográficas respectivas só estão disponíveis cerca de 4 meses depois do mês a que dizem respeito, assumindo, então, um carácter provisório, por ser significativo o número de registos feito muito depois do prazo estabelecido.

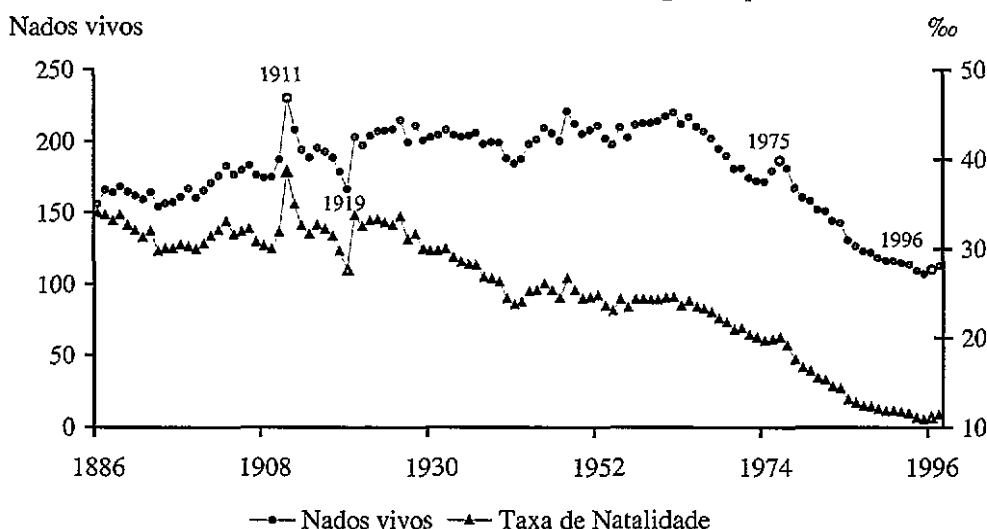
³ Ou seja, quando as estimativas da população residente são procuradas para alguns estudos, nomeadamente, o Inquérito ao Emprego.

2. Enquadramento Geral

2.1. A evolução da natalidade em Portugal desde 1886

O fenómeno da natalidade em Portugal, a partir de 1886, foi caracterizado por algumas oscilações bruscas alternadas com períodos de relativa estabilidade, até meados da década de 60, tendo, a partir daí, sofrido um quebra que, até ao ano de 1995, foi interrompida apenas nos dois anos seguintes à revolução de 1974. Desde 1965, exceptuando os anos de 1975 e 1976, o número de nados vivos, no nosso país, diminuiu, até que, em 1996, se deu um inesperado aumento — 2,94% — que se repetiu no ano seguinte — 2,44%.

Gráfico 1: Nados vivos e taxa de natalidade em Portugal no período de 1886 a 1997



Fonte: INE Anuários Demográficos, Estatísticas Demográficas e Séries de Estimativas Provisórias.

O aumento extraordinário no número dos nados vivos registados em 1911 (seguido de uma diminuição em 1912 para um nível ainda elevado) pode ser explicado pela introdução, nesse ano, da obrigatoriedade do registo civil (Carrilho, 1991). Ao número de crianças nascidas nesse ano pode estar erradamente acrescentada uma parcela referente àquelas que, tendo nascido nos anos anteriores, e por não ser obrigatório, não foram registadas oportunamente.

As baixas significativas observadas nos anos de 1918 e 1919, que se seguiram à diminuição devida à guerra e a um forte surto emigratório, reflectem os efeitos da gripe pneumónica (Carrilho, 1991), que terá provocado a morte de muitos jovens em idade de procriar.

Segundo Rosa (1996), a baixa dos nascimentos desde a década de 60 deve-se à estrutura populacional e ao nível de fecundidade: na década de 60, houve uma diminuição do peso das mulheres em idade fértil na população; nas duas décadas seguintes, o nível de fecundidade também foi importante. A autora apresenta como justificações para o declínio da fecundidade a alteração do calendário (o retardar da

idade ao primeiro filho leva a que cada mulher tenha menos filhos) e a nupcialidade. Nota, no entanto, que estas não são suficientes pois, nas décadas de 60 e 70, a fecundidade tornou-se mais precoce (só a partir de 80 começou a ser retardada), por outro lado, é cada vez mais significativa a proporção de nascimentos fora do casamento, o que desliga a natalidade cada vez mais da nupcialidade. O fenómeno sofre, portanto, influência de factores económicos e sociais ainda não identificados.

Os anos de 1975 e 1976 marcaram, como já foi referido, uma interrupção na tendência decrescente do número de nados vivos. Esse acréscimo ter-se-á devido à antecipação de casamentos e de nascimentos na sequência de perspectivas favoráveis em que se traduziram os aumentos de ordenados de 1974, 1975 e 1976 (Barata, 1985) e ao retorno de portugueses das ex-colónias (Carrilho, 1991).

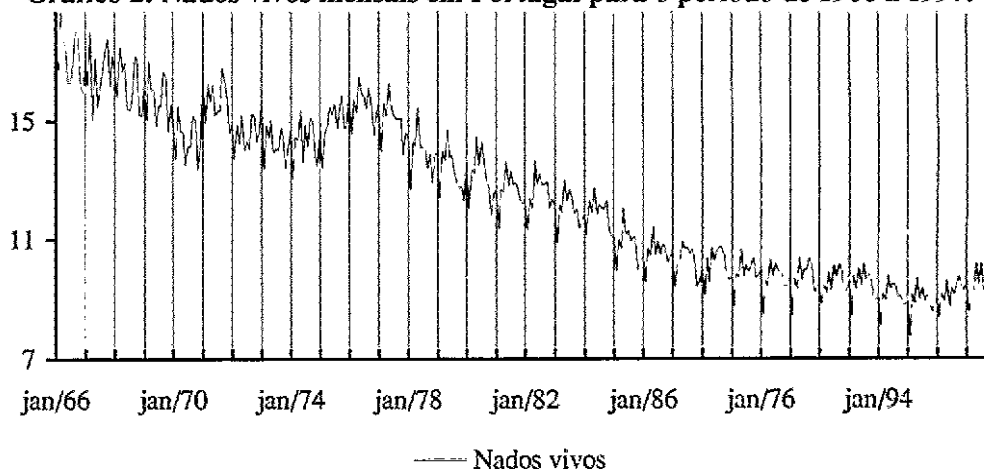
Se observarmos, no gráfico 1, os acréscimos da taxa de natalidade e número de nados vivos para os dois anos seguintes à revolução, notaremos que ambos foram significativos mas que o primeiro não foi tão importante. Isto leva-nos a concluir que o aumento dos nados vivos foi, em parte, devido ao aumento da população (provocado pelo referido retorno) mas não só, uma vez que também aumentou o número de nascimentos por cada mil habitantes. Então, houve, também, um maior desejo de procriar que o argumento de Barata justifica.

Carrilho e Peixoto (1993) destacam, como factores explicativos da queda da fecundidade, entre 1981 e 1992: o retardar do casamento e, conseqüentemente, da idade ao nascimento do primeiro filho; a difusão dos métodos contraceptivos; a dificuldade dos jovens no acesso à habitação e emprego; o prolongamento da escolaridade obrigatória; o maior grau de instrução e actividade profissional da mulher; a afirmação social e profissional do casal. Também aqui fica a ideia de que a natalidade tem de ser enquadrada no contexto socio-económico.

2.2. A sazonalidade nos nados vivos

A análise gráfica da série dos nados vivos mensais, sugere uma divisão desta

Gráfico 2: Nados vivos mensais em Portugal para o período de 1966 a 1997.



Fonte: INE Anuários Demográficos e Estatísticas Demográficas.

em dois períodos: aquele que vai de 1966 (ano a partir do qual existem dados) até cerca de 1975 e o período desde então até agora.

No primeiro período, não identificamos movimentos estritamente sazonais. Ao contrário, no segundo, observamos um padrão de sazonalidade dos nados vivos que é marcado por níveis mais elevados nos meses de Maio e Julho e mais baixos no mês de Fevereiro.

O facto de o aparecimento de sazonalidade nos nados vivos ter coincidido com a generalização dos métodos anticoncepcionais indica que esta será o reflexo do planeamento familiar. O fenómeno da sazonalidade estará, então, possivelmente, associado às preferências dos futuros pais quanto ao mês em que acontecerá o nascimento e, não tanto, a condições naturais.

2.3. A relação de masculinidade dos nados vivos

Chama-se relação de masculinidade à nascença ao número de nascimentos masculinos por cada cem nascimentos femininos:

$$\text{Relação de masculinidade à nascença} = \frac{\text{n.º de nados vivos de sexo masculino}}{\text{n.º de nados vivos de sexo feminino}} \times 100.$$

Este rácio é superior a 100 para quase todos os países para os quais existem dados, situando-se, para a maioria desses países, entre 104 e 107 (Shryock, 1976).

Esta relação é, por vezes, referida com uma constante demográfica igual a 105, devendo, segundo alguns autores, a observação de desvios acentuados indiciar falta de qualidade dos dados (Nazareth, 1996; Pressat, 1972). Refere Nazareth, "nos países com boa qualidade de dados a relação de masculinidade dos nascimentos anda à volta de 105, desde que se excluam as variações aleatórias sempre possíveis. A existência de desvios acentuados em relação a este valor médio não pode ser senão a consequência de erros observados, nomeadamente, omissões mais acentuadas num sexo do que noutro".

Se considerássemos o rácio acima definido constante e igual a 105, seríamos levados a concluir, pela observação dos seus valores desde o fim do século passado, que o registo dos nados vivos no nosso país é muito deficiente, pois esse valor foi sempre ultrapassado pela relação de masculinidade das crianças registadas, à excepção dos anos de 1911 a 1913 e 1928 em que foi ligeiramente inferior.

Julgamos que essa conclusão seria errada (pode dizer-se que existe qualidade destes dados pelo menos desde a década de 1980) pelo que essa "constante" não está de acordo com a realidade observada em Portugal.

A média dos valores observados no período de 1886 a 1997 foi de 106,6 e no período de 1980 até 1997 de 106,5. Quanto a este período, existe credibilidade das estatísticas dos nascimentos, pelo que julgamos que fará mais sentido tomar, pelo menos para o nosso país, como valor médio o de 106,5.

Quanto a considerá-lo constante, existe evidência de que isso pode não ser muito correcto (ainda em 1996 nasceram 108 rapazes para cada 100 raparigas).

2.4. A amostra

Existem dados mensais de nados vivos por sexo disponíveis a partir de 1967 mas os registos existentes até cerca do fim da década de 1970 carecem de credibilidade pelo que consideraremos como início da amostra o mês de Janeiro de 1980.

Os dados de que necessitamos estão disponíveis para o período até ao mês de Novembro de 1998, sendo os deste ano, ainda, provisórios. Porém, acreditamos que os números referentes aos meses do primeiro semestre não sofrerão mais alterações, razão pela qual os tomaremos como definitivos.

A amostra que utilizaremos para a estimação dos modelos de previsão é constituída pelas séries mensais de nados vivos masculinos e femininos ocorridos em Portugal (e cujos progenitores são residentes no nosso país) para o período de Janeiro de 1980 a Junho de 1998.

Consideraremos, a partir de agora,

$H_t = n^\circ$ de nados vivos de sexo masculino no período t ,

$M_t = n^\circ$ de nados vivos de sexo feminino no período t ,

$HM_t = H_t + M_t$,

e, ainda, a relação de masculinidade à nascença:

$$RMN_t = \frac{H_t}{M_t} \times 100.$$

3. Modelização ARIMA

3.1. O processo integrado misto sazonal: $SARIMA(p,d,q)(P,D,Q)_s$

X_t diz-se um processo misto integrado sazonal quando:

$$\phi_p(B)\Phi(B^s)(1-B)^d(1-B^s)^D X_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t,$$

em que o polinómio autoregressivo, $\phi_p(B)\Phi_p(B^s)$, é estacionário e o polinómio de médias móveis, $\theta_q(B)\Theta_Q(B^s)$, é invertível (ou seja, as suas raízes, em módulo, encontram-se fora do círculo unitário); ε_t é um ruído branco⁴.

⁴ $B^j X_t = X_{t-j}$.

S é o período da componente cíclica da série, que será igual a 12 no nosso caso.

Um processo estacionário $\{\varepsilon_t : t = 0, \pm 1, \pm 2, \dots\}$ diz-se ruído branco quando é formado por uma sucessão de variáveis aleatórias com a mesma distribuição, média constante $E[\varepsilon_t] = \mu_\varepsilon$, variância constante $V[\varepsilon_t] = \sigma_\varepsilon^2$ e covariância nula, $\gamma_k = 0$ para todo o inteiro $k \neq 0$ (Murteira e outros, 1993).

O modelo acima, designado abreviadamente por $SARIMA(p,d,q) \times (P,D,Q)_s$ é um processo não estacionário.

Tomando,

$$Y_t = (1-B)^d (1-B^s)^D X_t,$$

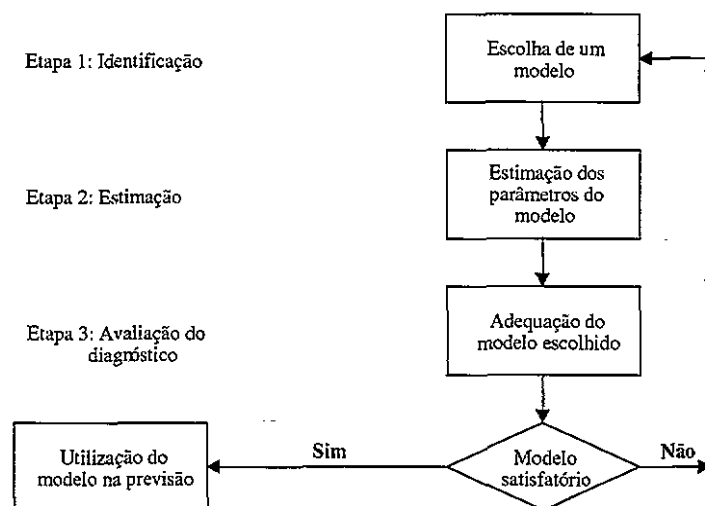
Temos o correspondente processo misto sazonal estacionário, habitualmente designado por $ARMA(p,q) \times (P,Q)_s$, definido por:

$$\phi_p(B)\Phi(B^s)Y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t.$$

3.2. Metodologia Box-Jenkins ⁵

Box e Jenkins, em 1970, propuseram uma metodologia de análise de uma série cronológica que tem por objectivo a obtenção de um modelo SARIMA que a descreva e, conseqüentemente, seja capaz de prever o seu comportamento futuro. Essa metodologia divide-se em três etapas e é, usualmente, esquematizada da seguinte forma:

Metodologia de Box-Jenkins



A etapa da identificação corresponde à escolha de um modelo SARIMA que, aparentemente, descreva a série em estudo e divide-se em:

I - Estacionarização da série e

II - Selecção de um modelo ARMA.

⁵Apresentamos um breve resumo desta metodologia, para maior desenvolvimento ver Box e outros (1994) e Murteira e outros (1993).

A primeira sub-etapa destina-se a, caso seja necessário, transformar a série de modo a estabilizar a variância, neutralizar a tendência e suprimir movimentos estritamente sazonais, ou seja, torná-la estacionária.

Segue-se a selecção de um modelo ARMA através da comparação das funções autocorrelação (FAC) e autocorrelação parcial (FACP) estimadas com os figurinos teóricos dessas funções para diferentes modelos⁶.

Na segunda etapa, procede-se à estimação dos parâmetros do modelo identificado.

O modelo estimado é submetido a uma avaliação do diagnóstico em termos da sua qualidade estatística (análise da qualidade estatística das estimativas sob determinados critérios) e da sua qualidade de ajustamento, ou seja, da sua adequação à série estudada (através de testes de hipóteses aos respectivos resíduos).

A não satisfação dos critérios de avaliação do diagnóstico por um modelo implica a sua rejeição. Tal como sugerido pelo algoritmo que descreve a metodologia, à rejeição de um modelo deverá seguir-se a escolha de outro (naturalmente orientada pela estimação e avaliação do/s anteriores). Caso contrário, é permitida a sua utilização na previsão.

No caso de o estudo da série levar a considerar aceitáveis vários modelos, deverá ser seleccionado um deles com base em critérios de qualidade de ajustamento e previsão. A avaliação da qualidade de previsão poderá ser feita através da simulação para um período para o qual são conhecidos os verdadeiros valores da série.

⁶Seja $\{X_t\}$ um processo estacionário, tem-se a sua função autocorrelação (FAC):

$$\rho_\tau = \frac{\text{cov}(X_t, X_{t+\tau})}{\text{Var}(X_t)}.$$

Considere-se o modelo de regressão de X_t sobre k dos seus valores passados, $k = 1, 2, \dots$:

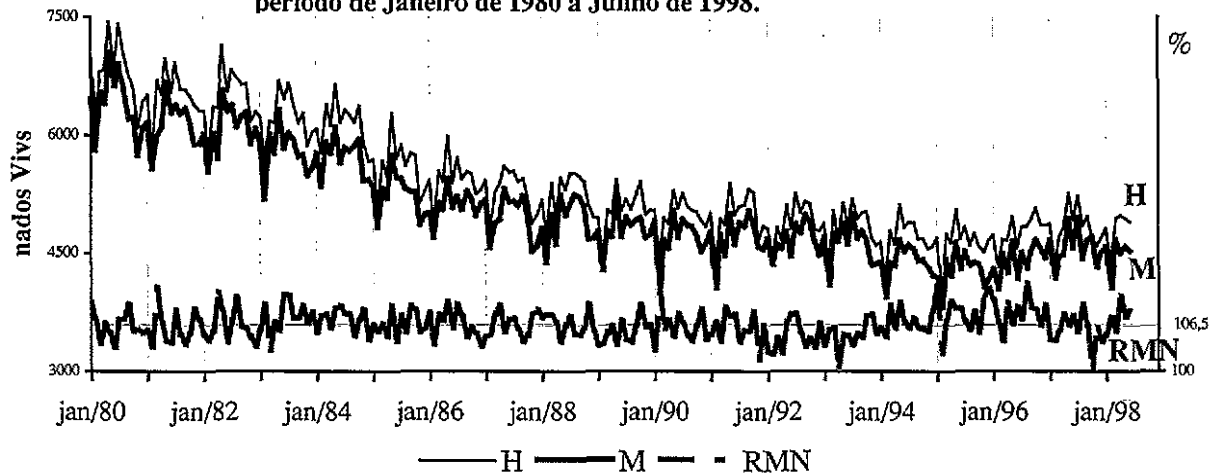
$$X_t = \phi_{k1}X_{t-1} + \phi_{k2}X_{t-2} + \dots + \phi_{kk}X_{t-k} + \varepsilon_t,$$

em que ε_t é um ruído branco de média nula. O coeficiente ϕ_{kk} , função de k , designa-se por autocorrelação parcial (FACP).

É possível o processo de identificação de modelos ARMA através da comparação referida pois as suas funções autocorrelação e autocorrelação parcial possuem características próprias.

3.3. Nados vivos e sua relação de masculinidade

Gráfico 5: Nados vivos masculinos e femininos e sua relação de masculinidade para o período de Janeiro de 1980 a Junho de 1998.



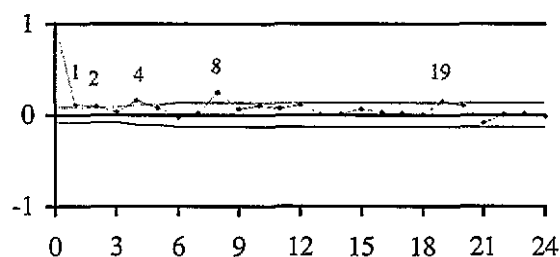
Fonte: INE Anuários Demográficos e Estatísticas Demográficas.

Como já foi referido, pretendemos prever o número de nascimentos de cada sexo. Por se tratarem de duas séries diferentes, em princípio, deveriam ser "tratadas" independentemente mas o seu comportamento, aparentemente e como seria de esperar, é semelhante.

Pela análise gráfica da série da relação de masculinidade à nascença, fica a ideia de que a sua variação em torno da constante 106,5 (valor médio da variável no período considerado) não tem um comportamento sistemático, ou seja, de que constitui um ruído branco.

Será, então, de testar a veracidade da hipótese de que a relação de masculinidade dos nados vivos é um ruído branco. Para isso, calcule-se a função de autocorrelação da série para os primeiros 24 desfasamentos⁷:

Gráfico 6: FAC estimada de RMN_t.



⁷As bandas apresentadas são : $\frac{1,25}{\sqrt{N}}$ para os três primeiros lags, $\frac{1,6}{\sqrt{N}}$ para os dois seguintes e

$\frac{2}{\sqrt{N}}$ para os restantes , sendo $N=n^{\circ}$ de observações, Murteira e outros (1993).

Como vemos, existe correlação significativa entre os valores da série e alguns valores desfasados, pelo que a variável não pode ser considerada um ruído branco.

Isto pode denunciar um comportamento diferente das séries de nascimentos dos dois sexos, justificando-se, portanto, que começemos por tratá-las em separado, podendo esperar-se a escolha de dois modelos distintos.⁸

Como podemos ver pela representação gráfica das séries (gráfico 5), estas não são estacionárias pois têm uma tendência decrescente e apresentam movimentos sazonais. Apesar de não identificarmos claramente um padrão de sazonalidade, pode dizer-se que os meses nos quais se observam os valores mais elevados são Maio e Julho, sendo os mais baixos observados no mês de Fevereiro.

Um dos factores que provoca a sazonalidade dos nascimentos pode ser o número de dias de cada mês. É legítimo pensar que nos meses mais longos ocorrem mais nascimentos, podendo argumentar-se a favor com os "casos" dos meses já referidos. Porém, este não será o único factor determinante da sazonalidade (por exemplo, em Agosto, nascem sempre menos crianças do que em Julho e ambos os meses têm 31 dias).

Podemos corrigir as séries do efeito do número de dias do mês, tomando, em vez do número de nados vivos, o número de nascimentos com vida por dia. Consideremos as variáveis:

$$\text{HPD}_t = \frac{H_t}{\text{Dias}_t}, \text{ n}^\circ \text{ de nados vivos de sexo masculino por dia no período } t,$$

$$\text{MPD}_t = \frac{M_t}{\text{Dias}_t}, \text{ n}^\circ \text{ de nados vivos de sexo feminino por dia no período } t,$$

com $\text{Dias}_t = \text{n}^\circ \text{ de dias do período (mês) } t$.

O mês de Fevereiro já não é o que apresenta os valores mais baixos em todos os anos (sendo os mínimos anuais atingidos, por vezes, em Novembro e Março). O pico de Julho parece justificar-se pelo facto de este mês ter 31 dias, pois já não acontece nas variáveis corrigidas, surgindo o mês de Setembro como um daqueles em que acontecem mais nascimentos com vida por dia. Não é, agora, tão claro um padrão de sazonalidade.

Com vista à obtenção do melhor modelo de previsão para os nados vivos aplicaremos à série HMPD^9 a metodologia *Box-Jenkins*.

3.5. Modelização da série do total de nados vivos

A análise da série de relação de masculinidade à nascença justificou que se começasse por tratar as duas variáveis que relaciona em separado. Porém, essa análise

⁸ Foi o que fizemos tendo chegado a dois modelos muito parecidos. Comparámos os resultados obtidos através da simulação da previsão para períodos passados, e concluímos que não há vantagem em estimar um modelo para cada sexo. Optámos, então, por estimar apenas um modelo e fazer previsões para o número total de nados vivos que será repartido por sexos segundo a relação de masculinidade de 106,5.

⁹ Foi também modelizada a série HM mas os resultados obtidos com as série corrigida foram melhores.

levou a que se optasse por ajustar às duas séries um modelo $SARIMA(0,1,1)(0,1,1)_{12}$, cujos parâmetros estimados não diferem muito. Neste caso, será que se justifica a estimação de dois modelos diferentes, tendo sido verificado que o comportamento das duas séries é semelhante?

Já vimos que existe uma relação entre as duas séries que, apesar de não ser constante, varia pouco em torno de um valor médio de 106,5 nascimentos masculinos por cada 100 nascimentos femininos. Em alternativa à construção de modelos diferentes para os nascimentos de sexos diferentes, poderíamos estimar um só modelo para o número total de nascimentos. As estimativas obtidas para esse total seriam, posteriormente, repartidas tomando a relação de masculinidade à nascença igual ao valor médio referido.

3.5.1. O modelo estimado

Ajustando um modelo $SARIMA(0,1,1)(0,1,1)_{12}$ à série dos nados vivos (considerando o intervalo de estimação de Janeiro de 1980 a Junho de 1998), obtiveram-se as seguintes estimativas:

$$(1-B)(1-B^{12})HMPD_t = (1 - \underset{t_{\theta_1}=6,262}{0,3986B})(1 - \underset{t_{\theta_1}=69,86}{0,9198B^{12}})\varepsilon_t,$$

onde ε_t é um ruído branco.

Também aqui, a estimativa do parâmetro do factor sazonal de médias móveis não se afasta muito de 1 mas, de novo através de um teste de hipóteses, concluímos que diferem significativamente, portanto, o respectivo polinómio é invertível.

4. Modelização estrutural

4.1. Introdução

Os modelos estruturais de séries cronológicas foram desenvolvidos por Andrew Harvey, em 1984 e 1985, e, ao contrário dos modelos SARIMA, descrevem a série cronológica directamente em termos das suas componentes tradicionais: a tendência, o ciclo, a sazonalidade e, ainda, a componente residual.

Harvey (1997) apresenta duas grandes objecções à abordagem ARIMA: a primeira diz respeito à metodologia de selecção de um modelo estando a segunda relacionada com a própria classe de modelos.

Quanto à primeira objecção, o autor defende que o procedimento não funciona. No seu livro, Box e Jenkins (1976), referiram que os métodos propostos se destinavam a modelos simples para grandes amostras. De facto, com centenas de observações, é possível identificar através do correlograma de uma série um $AR(1)$,

ou MA(2), ou mesmo um ARMA(1,1) mas o mesmo não se pode dizer de modelos mais complexos, e para amostras mais pequenas. Harvey nota que, mesmo reconhecendo a função autocorrelação teórica de, por exemplo, um processo ARMA(3,2), é quase impossível alguém identificá-lo através da função autocorrelação estimada.

Por outro lado, defende que, se a série não é estacionária, sendo, por isso, a análise baseada nas suas diferenças, a selecção de um modelo inadequado pode ter consequências graves ao nível da previsão. Na sua opinião, a utilização de critérios como o AIC (Akaike Information Criterion) para ultrapassar as dificuldades na selecção de um modelo traz poucos benefícios, podendo, até, ser prejudicial.

A segunda objecção em relação à modelização ARIMA diz respeito à classe de modelos. O autor não concorda com a opinião de alguns estudiosos, segundo a qual existe um teorema que diz que todas as séries são geradas por um processo daquele tipo, bastando procurar para encontrar o correcto. Aquela classe compreende muito modelos que não têm uma interpretação clara (e cujas funções de previsão podem ter propriedades indesejáveis) que podem ser os escolhidos, dadas as dificuldades na selecção.

A perspectiva tradicional da análise de séries cronológicas pressupõe que o conjunto de observações que constitui a série é o resultado da interacção de componentes elementares como a tendência, ciclo e sazonalidade. A componente residual, ou aleatória, é constituída por movimentos irregulares, não previsíveis, ou seja, inclui tudo o que não é sistemático. Esta abordagem é a chamada modelização de componentes, ou análise de decomposição e assume que as componentes são determinísticas¹⁰.

Uma vez que uma tendência determinística é muito restritiva, a solução óbvia é torná-la mais flexível, deixando os parâmetros do nível e do declive variar ao longo do tempo (Harvey, 1997).

Num modelo estrutural, assume-se que estes parâmetros são passeios aleatórios¹¹ o que leva a uma tendência aleatória em que o nível e o declive podem evoluir ao longo do tempo. As previsões desse modelo dão mais importância às observações mais recentes, quanto mais depressa o nível e o declive mudam, mais são desprezadas as observações passadas. A tendência determinística é o caso limite em que os hiperparâmetros que permitem que as suas componentes (nível e declive) variem são iguais a zero¹².

À tendência e componente irregular podem ser acrescentadas outras componentes como o ciclo e a sazonalidade que são aleatórias (podendo, também, no limite, tornar-se determinísticas).

¹⁰ Os interessados podem encontrar a apresentação dos modelos de decomposição em Murteira e outros (1993).

¹¹ Um processo estocástico $\{X_t : t = 0, \pm 1, \pm 2, \pm 3, \dots\}$ diz-se um passeio aleatório quando, para cada

$t = 0, \pm 1, \pm 2, \pm 3, \dots$, se tem $X_t = X_{t-1} + \varepsilon_t$, em que ε_t é um ruído branco.

¹² Estes hiperparâmetros são as variâncias das perturbações aleatórias das componentes. Isto tornar-se-á mais claro com a formalização do modelo que apresentaremos de seguida.

Os efeitos sazonais dizem respeito ao longo prazo, pois o padrão sazonal é projectado indefinidamente no futuro, devendo ser a estimativa do último, no fim da série, uma vez que estes padrões podem variar ao longo do tempo. Ora, é isto que faz o modelo estrutural, ao admitir a aleatoriedade da componente sazonal.

Na opinião de Harvey (1997), a formalização da sazonalidade de forma determinística, através de variáveis "dummy", não é aconselhável pois acontecem, frequentemente, alterações nos padrões sazonais das séries. O argumento de alguns econométricos de que os efeitos sazonais devem ser considerados estacionários, pois o "Inverno não pode tornar-se Verão", é refutado pelo autor, que refere que existem muitos exemplos em que estes efeitos variam.

O modelo estrutural é formalizado em termos das suas componentes, que têm uma interpretação directa. Em termos econométricos convencionais, pode pensar-se num modelo de regressão em que as variáveis explicativas são funções do tempo e em que os parâmetros variam com o tempo (Harvey, 1993)

As previsões num modelo estrutural são construídas automaticamente pelo filtro de Kalman, processo de estimação recursiva, ótimo no sentido clássico de minimização do erro quadrático médio. A tendência é extraída por um algoritmo de alisamento e os parâmetros que determinam a sua evolução são estimados pelo método da máxima verosimilhança, usando, de novo, o filtro de Kalman.

A análise tradicional de séries cronológicas salienta o papel da diferenciação na construção de modelos de séries não estacionárias; apesar de os modelos estruturais poderem ser representados em diferenças, nem sempre é necessário fazê-lo para obter um modelo adequado, podendo estes ser representados em níveis, o que torna a sua interpretação mais clara.

A modelização estrutural pode ser implementada através do package STAMP¹³. Não sendo necessário conhecer o filtro de Kalman para usar este package, o utilizador comum pode concentrar-se apenas na selecção de um modelo adequado e sua interpretação (Harvey, 1997).

4.2. Formalização do modelo

O modelo é formalizado do seguinte modo:

$$y_t = \mu_t + \gamma_t + \psi_t + v_t + \varepsilon_t, \text{ com } \varepsilon_t \sim \text{wn}(0, \sigma_\varepsilon^2).$$

A equação acima relaciona as observações da variável dependente com as componentes estruturais, respectivamente, a tendência, a sazonalidade, o ciclo, uma componente autoregressiva de primeira ordem e a componente irregular. A primeira parcela corresponde à tendência que, por sua vez, é formalizada através das duas equações:

¹³ A última versão do STAMP (Structural Time Series Analyser, Modeller and Predictor) foi construída em 1995 por Siem Koopman, Andrew Harvey.

$$\begin{aligned}\mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t, \text{ com } \eta_t \sim \text{wn}(0, \sigma_\eta^2) \text{ e} \\ \beta_t &= \beta_{t-1} + \zeta_t, \text{ com } \zeta_t \sim \text{wn}(0, \sigma_\zeta^2),\end{aligned}$$

onde β_t é o declive da tendência μ_t . A componente irregular, ε_t , as perturbações aleatórias do nível, η_t , e do declive, ζ_t , admitem-se não correlacionadas entre si. A componente declive pode, quando apropriado, ser retirada do modelo ou, ainda, ser definida como um processo autoregressivo de primeira ordem, chamando-se tendência amortecida:

$$\beta_t = \rho_\beta \beta_{t-1} + \zeta_t, \text{ com } \zeta_t \sim \text{wn}(0, \sigma_\zeta^2) \text{ e } 0 < \rho_\beta < 1.$$

A sazonalidade pode ser definida na forma tradicional:

$$\gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t, \text{ com } \omega_t \sim \text{wn}(0, \sigma_\omega^2),$$

onde s é o período da sazonalidade da série, ou então na forma trigonométrica:

$$\gamma_t = \sum_{j=1}^{\lfloor s/2 \rfloor} \gamma_{j,t},$$

onde cada $\gamma_{j,t}$ é gerado por:

$$\begin{bmatrix} \gamma_{j,t} \\ \gamma_{j,t}^* \end{bmatrix} = \begin{bmatrix} \cos \lambda_j & \text{sen} \lambda_j \\ -\text{sen} \lambda_j & \cos \lambda_j \end{bmatrix} \begin{bmatrix} \gamma_{j,t-1} \\ \gamma_{j,t-1}^* \end{bmatrix} + \begin{bmatrix} \omega_{j,t} \\ \omega_{j,t}^* \end{bmatrix}, \quad j=1, \dots, \lfloor s/2 \rfloor, \\ t=1, \dots, T,$$

sendo $\lambda_j = \frac{2\pi j}{s}$ a frequência, em radianos, e as perturbações aleatórias, $\omega_{j,t}$ e $\omega_{j,t}^*$, ruídos brancos não correlacionados entre si, de média nula e desvio padrão comum σ_ω .

A especificação estatística do ciclo, ψ_t é dada por:

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho_\psi \begin{bmatrix} \cos \lambda_c & \text{sen} \lambda_c \\ -\text{sen} \lambda_c & \cos \lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix}, \quad t=1, \dots, T,$$

onde $0 < \rho_\psi \leq 1$ e $0 \leq \lambda_c \leq \pi$ são, respectivamente, o factor de amortecimento e a frequência do ciclo, em radianos; κ_t e κ_t^* são ruídos brancos não correlacionados entre si, de média nula e variância comum σ_κ^2 . O período do ciclo é igual a $\frac{2\pi}{\lambda_c}$.

A componente autoregressiva é definida da seguinte forma:

$$v_t = \rho_v v_{t-1} + \xi_t, \text{ com } \xi_t \sim \text{wn}(0, \sigma_\xi^2) \text{ e } 0 < \rho_v < 1.$$

As perturbações aleatórias de cada componente são não correlacionadas entre si.

As variâncias das perturbações aleatórias das componentes, juntamente com as frequências e factores de amortecimento dos ciclos (e declive) e os coeficientes da componente autoregressiva constituem os hiperparâmetros dos modelos estruturais. Cada uma das componentes do modelo torna-se determinística caso o hiperparâmetro que permite que varie seja igual a zero.

Se admitirmos a normalidade das perturbações do modelo, estes hiperparâmetros podem ser estimados pelo método da máxima verosimilhança, o que pode ser feito no domínio tempo, através do filtro de Kalman, ou no domínio frequência. Feita esta estimação, utiliza-se o modelo no espaço de estados para fazer previsões e construir estimadores das suas componentes não observáveis.

4.3. Modelos em espaço de estados

O tratamento estatístico dos modelos estruturais baseia-se na sua representação em termos de espaço de estados, por forma a que a estimação das suas componentes possa ser feita com o filtro de Kalman. De seguida, apresentaremos uma formalização simples desta representação.

Seja $y_t, t=1, 2, \dots$, um sistema dinâmico que evolui ao longo do tempo e que pode, por hipótese, em cada momento, ser caracterizado pelos valores de m variáveis (as quais podem não ser directamente observáveis). Estas m variáveis, designadas por variáveis de estado, podem formalizar-se no seguinte vector:

$$\alpha_t = [\alpha_{1t} \quad \alpha_{2t} \quad \dots \quad \alpha_{mt}]^T,$$

que dá, para cada t , o estado do processo.

A equação das observações, ou de medida:

$$y_t = Z_t \alpha_t + \varepsilon_t, t=1, 2, \dots, T,$$

onde Z_t é um vector ($1 \times m$) e ε_t um ruído branco de média nula e variância igual a $\sigma^2 h$, relaciona as observações, y_t , e o vector de estados, α_t .

A equação de transição:

$$\alpha_t = T_t \alpha_{t-1} + \pi_t$$

caracteriza a evolução no tempo do vector de estados, ainda que este possa não ser directamente observável. Na equação acima tem-se T_t , matriz ($m \times m$), e π_t vector aleatório ($m \times 1$) com $V[\pi_t] = \sigma^2 Q_{(m \times m)}$. Os elementos de π_t são ruídos brancos que se admitem não correlacionados entre si nem com ε_t .

O conjunto das equações das observações e de transição define um modelo em espaço de estados.

A representação acima admite que as matrizes Z_t e T_t podem variar com o tempo. Porém, no caso dos modelos estruturais, elas são invariantes com t , tomando-se, então, no seu lugar, Z e T .

4.3.1. Modelos estruturais em espaço de estados

A representação de modelos estruturais em espaço de estados é quase imediata. Considere-se, por exemplo, o chamado modelo estrutural básico que inclui as componentes tendência, sazonalidade e residual:

$$\begin{aligned} y_t &= \mu_t + \gamma_t + \varepsilon_t, \\ \mu_t &= \mu_{t-1} + \beta_{t-1} + \eta_t, \\ \beta_t &= \beta_{t-1} + \zeta_t, \\ \gamma_t &= -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t, \end{aligned}$$

com $\varepsilon_t, \eta_t, \zeta_t$ e ω_t ruídos brancos de média nula, não correlacionados entre si e cujas variâncias são, respectivamente, $\sigma_\varepsilon^2, \sigma_\eta^2, \sigma_\zeta^2, \sigma_\omega^2$.

A representação em espaço de estados do modelo estrutural básico, para o caso particular de dados trimestrais, é:

$$\left\{ \begin{array}{l} y_t = \underbrace{\begin{bmatrix} 1 & 0 & 1 & 0 & 0 \end{bmatrix}}_Z \underbrace{\begin{bmatrix} \mu_t \\ \beta_t \\ \gamma_t \\ \gamma_{t-1} \\ \gamma_{t-2} \end{bmatrix}}_{\alpha_t} + \varepsilon_t, \\ \alpha_t = \underbrace{\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}}_T \alpha_{t-1} + \underbrace{\begin{bmatrix} \eta_t \\ \zeta_t \\ \omega_t \\ 0 \\ 0 \end{bmatrix}}_{\pi_t} \end{array} \right.$$

O vector de estados inclui o número mínimo de componentes necessário para caracterizar o processo no momento t : nível, tendência e as três mais recentes estimativas para os factores sazonais.

4.4. Estimação recursiva e o filtro de Kalman

O filtro de Kalman desempenha o mesmo papel para os modelos de séries cronológicas em espaço de estados que os métodos de mínimos quadrados para os modelos de regressão (Koopman e outros, 1995).

Num modelo em espaço de estados, o vector de estados é essencial para a caracterização do sistema pelo que, sendo este, normalmente, desconhecido, é necessário estimá-lo; isto pode ser feito através do método de estimação recursiva chamado filtro de Kalman, que é óptimo no sentido de proporcionar os estimadores de menor erro quadrático médio (EQM). De seguida, explicitaremos a sua forma para os modelos acima definidos apresentada em Costa (1995)¹⁴.

4.4.1. O filtro de Kalman

Admitamos conhecidos os hiperparâmetros do modelo. Sejam, em t-1:

a_{t-1} , o estimador (ou estimativa) de menor EQM de α_{t-1} , baseado na informação disponível até t-1, inclusive, e

$\sigma^2 P_{t-1}$, a matriz do erro quadrático médio desse estimador.

Dados a_{t-1} e $\sigma^2 P_{t-1}$, temos as equações de previsão:

$$a_{t|t-1} = T a_{t-1} \text{ e}$$

$$P_{t|t-1} = T P_{t-1} T^T + Q,$$

em que $a_{t|t-1}$ é o estimador (previsor) de menor EQM de α_t em t-1 de matriz de covariâncias $\sigma^2 P_{t|t-1}$.

A previsão óptima, ou seja, de menor EQM, de y_t em t-1 é:

$$\hat{y}_{t|t-1} = Z a_{t|t-1}.$$

Conhecido y_t , procede-se à actualização das estimativas do vector de estados e respectiva matriz do erro quadrático médio, através das equações de actualização:

$$a_t = a_{t|t-1} + K_t e_t \text{ e}$$

$$P_t = (I - K_t Z) P_{t|t-1},$$

¹⁴ Existem formas mais gerais, por exemplo, para o caso de introdução no modelo de variáveis explicativas e de intervenção que também é possível na modelização estrutural mas que não foi, aqui, considerado. Uma forma mais geral, pode ser encontrada em Koopman e outros (1995), Harvey(1993).

onde $K_t = P_{t|t-1} Z^T [Z P_{t|t-1} Z^T + h]^{-1}$ e $e_t = y_t - \hat{y}_{t|t-1}$ são, respectivamente, o chamado ganho de Kalman e o erro de previsão a um passo, ou inovação.

O conjunto das equações de previsão e actualização constitui o filtro de Kalman.

4.4.2. Previsão

O estimador, ou estimativa, de menor erro quadrático médio do vector de estados em $T+h$, no momento T (ou seja, T e h são, respectivamente, a origem e o horizonte de previsão) é:

$$a_{T+h|T} = T a_{T+h-1|T}, h = 1, 2, \dots,$$

tendo-se, no seu erro de previsão,

$$P_{T+h|T} = T P_{T+h-1|T} T^T + Q, h = 1, 2, \dots$$

Daqui se obtém o predictor das observações, y_{t+h} :

$$\hat{y}_{T+h|T} = Z a_{T+h|T},$$

de variância,

$$V[\hat{y}_{T+h|T}] = \sigma^2 (Z P_{T+h|T} Z^T + h).$$

4.4.3. Alisamento e decomposição

Quando, para além do objectivo da previsão, existe também o de decomposição da série, tendo em vista a obtenção das melhores estimativas possíveis das componentes, procede-se ao alisamento em que se estima cada vector de estados tendo em conta, não só as observações passadas, como na filtragem, mas a informação de toda a amostra.

Por ser o nosso objectivo principal a previsão, apresentamos, apenas, uma formalização simples deste procedimento.

O alisamento consiste na aplicação dos princípios da estimação recursiva em ordem inversa à das observações, tendo como valores iniciais as estimativas a_T e P_T obtidas pelo filtro de Kalman.

As equações de alisamento são:

$$a_{t|T} = a_t + P_t^* (a_{t+1|T} - T a_t) \text{ e}$$

$$P_{t|T} = P_t + P_t^* (P_{t+1|T} - P_{t+1|t}) P_t^{*T},$$

com $P_t^* = P_t T_{t+1}^T P_{t+1|t}^{-1}$, $t = T-1, T-2, \dots, 1$.

4.5. Ajustamento dos modelos estruturais

Ainda que diferente da primeira fase de identificação da modelização ARIMA, na modelização estrutural, há que fazer uma análise da série em estudo por forma a escolher qual o modelo estrutural que será mais adequado, ou seja, que componentes devem ser incluídas e se algumas devem ser consideradas, à partida, determinísticas.

Há que ter em conta, também, a eventual necessidade de efectuar alguma transformação não linear sobre a série de modo a garantir a adequação da formalização aditiva do modelo.

O filtro de Kalman assume o conhecimento dos hiperparâmetros, que devem ser, previamente, estimados. Usualmente, são-no através de métodos de máxima verosimilhança em que se supõe a normalidade das perturbações aleatórias das componentes¹⁵.

Após este processo de estimação, donde poderão resultar estimativas nulas para as variâncias das perturbações, o que indicará que as respectivas componentes são determinísticas, a série será filtrada uma última vez. Desta filtragem, resultarão os erros de previsão (resíduos) a partir dos quais serão calculadas as estatísticas a utilizar na avaliação do modelo ajustado bem como a estimativa do estado final, que contém a informação necessária para efectuar previsões.

Após a estimação do modelo, pode proceder-se a uma avaliação de diagnóstico, através de testes à normalidade dos resíduos, heterocedasticidade e aleatoriedade bem como algumas estatísticas indicadoras da qualidade de ajustamento. Os testes e indicadores de qualidade de ajustamento utilizados para os modelos estruturais (e fornecidos pelo STAMP) são os comumente utilizados.¹⁶

4.6. Um modelo para a série dos nados vivos¹⁷

Com o mesmo objectivo que nos guiou na secção anterior, o de obter boas previsões para o número de nascimentos de ambos os sexos ocorridos por mês, experimentaremos a modelização estrutural descrita atrás.

Testaremos se, para a nossa série, é verdade o que Harvey defende, ou seja, que modelização estrutural conduz a melhores preditores do que os alcançados com a modelização ARIMA.

¹⁵ Para uma consulta dos métodos utilizados pelo STAMP, veja-se Koopman e outros (1995).

¹⁶ Ver Koopman e outros (1995) e Costa (1995).

¹⁷ Tendo concluído que não há vantagem em estimar dois modelos diferentes para os dois sexos, estimaremos apenas um modelo para o total de nados vivos; tomaremos esta série corrigida do número de dias do mês também por havermos constatado a vantagem de fazê-lo.

4.6.1. O modelo estimado

A modelização estrutural da série dos nados vivos conduziu-nos ao seguinte modelo:

$$\text{HMPD}_t = \mu_t + \gamma_t + v_t + \varepsilon_t, \text{ com } \varepsilon_t \sim \text{wn}(0, \sigma_\varepsilon^2 = 8,129),$$

em que a tendência e as componentes sazonal e autoregressiva de primeira ordem são definidas, respectivamente, por:

$$\mu_t = \mu_{t-1} + \eta_t, \text{ com } \eta_t \sim \text{wn}(0, \sigma_\eta^2 = 13,67),$$

$$\gamma_t = -\sum_{j=1}^{11} \gamma_{t-j} + \omega_t, \text{ com } \omega_t \sim \text{wn}(0, \sigma_\omega^2 = 0,1335),$$

$$v_t = 0,5755v_{t-1} + \xi_t, \text{ com } \xi_t \sim \text{wn}(0, \sigma_\xi^2 = 22,18).$$

Em espaço de estados, temos:

$$\left. \begin{array}{l} \text{HMPD}_t = \underbrace{[1 \ 1 \ 0 \ 0 \ \dots \ 0 \ 1]}_Z \underbrace{\begin{bmatrix} \mu_t \\ \gamma_t \\ \gamma_{t-1} \\ \gamma_{t-2} \\ \vdots \\ \gamma_{t-10} \\ v_t \end{bmatrix}}_{\alpha_t} + \varepsilon_t, \\ \alpha_t = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 & \dots & -1 & -1 & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ & & & & \vdots & & & \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0,5755 \end{bmatrix} \alpha_{t-1} + \begin{bmatrix} \varepsilon_t \\ \omega_t \\ 0 \\ 0 \\ \vdots \\ 0 \\ \xi_t \end{bmatrix} \\ \underbrace{\hspace{10em}}_{\hat{T}} \underbrace{\hspace{1em}}_{\pi_t} \end{array} \right\}$$

A estimativa para o estado final, a partir do qual serão calculadas as previsões é:

$$a_{\text{Jun}/98} = \begin{bmatrix} 306,7 \\ 6,963 \\ 15,70 \\ -0,8867 \\ 6,192 \\ -13,21 \\ -8,556 \\ -14,47 \\ -6,797 \\ 1,256 \\ 15,13 \\ -0,1736 \\ 0,9956 \end{bmatrix}$$

O último padrão de sazonalidade estimado e que será projectado no futuro é:

Janeiro	Fevereiro	Março	Abril	Maio	Junho
-8,556	-13,21	-6,192	-0,8867	15,70	6,963
Julho	Agosto	Setembro	Outubro	Novembro	Dezembro
11,24	-0,1736	15,13	1,256	-6,797	-14,47

5. Modelo Econométrico

A influência de vários factores sócio-económicos sobre o fenómeno da natalidade tem sido referida por vários autores. O objectivo desta secção é a construção de um modelo econométrico que ponha em evidência a influência de alguns desses factores e que permita a melhoria das previsões em relação aos modelos univariados estudados atrás. Os modelos univariados já mostraram ser capazes de fornecer boas previsões mas sobretudo de variáveis "bem comportadas", ou seja, sem inversões de tendência como a ocorrida com o número de nados vivos em Portugal no ano de 1996. É de esperar que um modelo que faça depender uma variável apenas das suas observações passadas preveja de acordo com a tendência verificada até então.

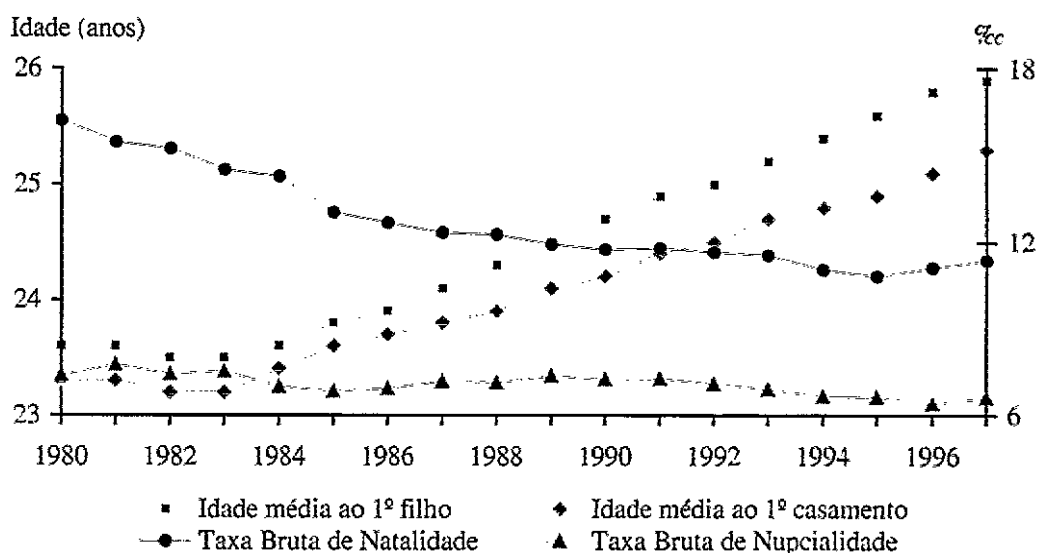
Acreditamos que a conjuntura económica do país influi, ao nível do casal, no desejo de ter ou não ter filhos o que, ao nível do país, irá reflectir-se no número dos nascimentos.

O fenómeno da nupcialidade surge, naturalmente, associado ao da natalidade. O declínio da nupcialidade (e o atraso da idade ao primeiro casamento) foi acompanhado pelo aumento da idade com que as mulheres têm o seu primeiro filho (diminuindo o

número de filhos que cada mulher tem ao longo da sua vida) e, conseqüentemente, o decréscimo do número total de nascimentos.

Porém, atente-se no facto de que o aumento da taxa bruta de natalidade no ano de 1996 não foi precedido por um aumento da taxa bruta de nupcialidade, ao contrário, esta sofreu um acréscimo apenas no ano seguinte¹⁸.

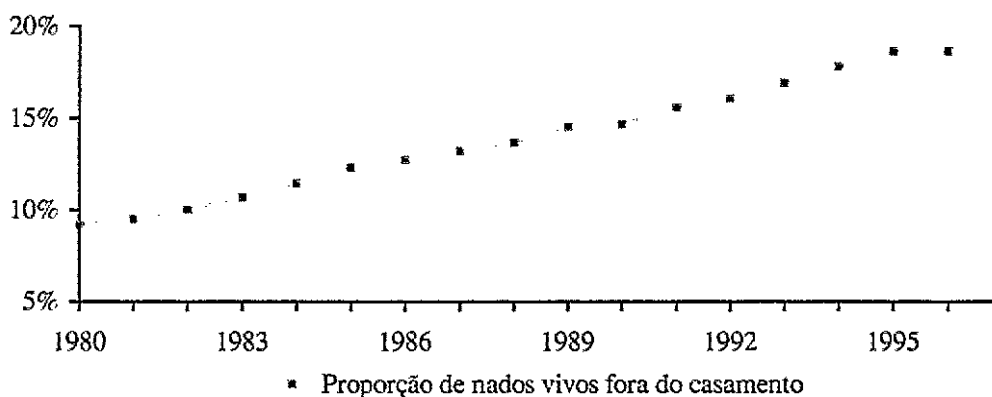
Gráfico 18: Taxas brutas de natalidade e de nupcialidade e idades médias da mulher ao primeiro filho e ao primeiro casamento, em Portugal, de 1980 a 1997.



Fonte: INE Séries de Estimativas Provisórias.

Não podemos encarar o número de casamentos como único factor determinante do número dos nascimentos pois a proporção destes que ocorrem fora dos casamentos é, hoje em dia e cada vez mais, muito significativa. No ano de 1997, o número de crianças nascidas cujos pais não eram casados representou 19,5% do total de nascimentos, no nosso país.

Gráfico 19: Proporção de nados vivos fora do casamento, em Portugal de 1980 a 1997



Fonte: INE Estimativas Demográficas.

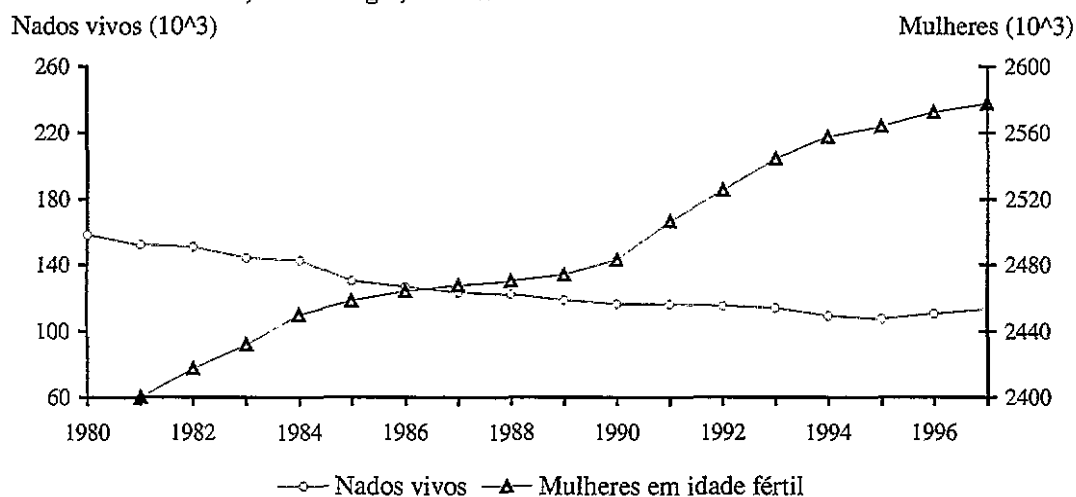
¹⁸ Taxa Bruta de Nupcialidade = $\frac{\text{número de casamentos celebrados}}{\text{População média}} \times 1000$, (Pressat, 1972).

Portanto, não haverá, entre os dois fenómenos, a relação de causalidade que, porventura, já existiu mas não é de excluir a hipótese de que estão correlacionados por sofrerem influências comuns. É legítimo pensar que, num dado momento, uma boa situação económica e expectativas de prosperidade de um país podem levar a que haja mais casais a optarem pelo casamento e, também, mais casais a decidirem ter um filho.

O número de nascimentos ocorridos num período pode ser considerado função da estrutura da população, em particular do número de mulheres em idade de procriar (considera-se, habitualmente, como idades férteis das mulheres as que vão dos 15 aos 49 anos). A ser assim, observaríamos variações destes dois valores no mesmo sentido.

A hipótese de que a evolução desfavorável da estrutura da população, no sentido referido acima, provocou o declínio do fenómeno em estudo desde a década de oitenta até ao ano de 1995 pode ser facilmente refutada pela observação do gráfico seguinte.

Gráfico 20: Mulheres com idades compreendidas entre os 15 e os 49 anos e nados vivos anuais, em Portugal, de 1981 a 1997



Fonte: INE Estatísticas Demográficas e Séries de Estimativas Provisória

Ao contrário, a população feminina susceptível de procriar tem vindo a aumentar. Não fará, portanto, sentido fazer depender desta variável a que pretendemos explicar.

Na fase de construção do modelo econométrico explicativo do número de nados vivos mensais, estudaremos a influência de dois indicadores económicos passíveis de reflectir a conjuntura económica que, por sua vez, será relevante na decisão de ter um filho: o índice de produção industrial da indústria transformadora e o valor das vendas de gasolina. Em relação a estas variáveis, tomaremos os seus desfasamentos de nove e mais meses, de modo a que representem, para cada mês, o período em que as decisões de ter os filhos aí nascidos foram tomadas.

Será, também, em conformidade com o que foi dito acerca da possível correlação entre os fenómenos da natalidade e da nupcialidade, estudada a

significância do número de casamentos celebrados cerca de nove meses antes na explicação do número de nascimentos ocorridos num dado período.

Mais uma vez, não esqueceremos a influência que o número de dias de um mês pode ter sobre os nascimentos ali ocorridos que, desta feita, será estudada directamente.

A sazonalidade será modelizada através de variáveis "dummy" mensais.

Apesar da inclusão no modelo de outras variáveis explicativas, não desprezaremos a possível influência dos valores passados da própria variável dependente pelo que incluiremos, também, alguns seus desfasamentos.

5.1. O Modelo

5.1.1. As variáveis do modelo

A estimação de vários modelos alternativos com mais ou menos desfasamentos das variáveis explicativas, mais ou menos observações passadas da variável dependente levou à escolha de um modelo com as seguintes variáveis:

- Variável dependente

ΔHM_t = número de nados vivos no período t (em primeiras diferenças).

- Variáveis explicativas

ΔHM_{t-i} = variável dependente desfasada de i períodos, $i = 1, 9$,

$\Delta \text{Transf}_{t,i}$ = índice de produção industrial da indústria transformadora em primeiras diferenças e desfasado de i períodos, $i = 9, 10$,

$\Delta \text{Gasoli}_{t,9}$ = valor das vendas de gasolina em primeiras diferenças e desfasado de nove meses,

ΔDias_t = número de dias do período (mês) t, em primeiras diferenças,

S_{it} = variável que toma o valor 1 para os primeiros seis meses de cada ano e zero para os restantes,

M_{it} = variável "dummy" mensal que toma o valor 1 para o mês i e zero para os restantes, $i = 5, 6, 8, 9, 12$.

Foram, inicialmente, incluídas variáveis explicativas referentes aos casamentos (desfasamentos de nove a doze meses do número de casamentos em primeiras diferenças). Contudo, obtiveram-se, para todas, coeficientes estatisticamente nulos o que ditou a sua exclusão do modelo.

Daí, concluímos que não temos no número de casamentos um indicador avançado para o número de nascimentos; confirma-se a ideia de que a natalidade, hoje em dia, já não é influenciada pela nupcialidade.

5.1.2. O modelo estimado

O modelo estimado para a amostra até Junho de 1998 é o seguinte (entre parênteses encontram-se as estatísticas t das estimativas dos parâmetros):

$$\begin{aligned} \Delta HM_t = & -250,3 - 0,2786 \Delta HM_{t-1} + 0,1640 \Delta HM_{t-9} + 15,18 \Delta \text{Transf}_{t-9} + \\ & + 7,097 \Delta \text{Transf}_{t-10} + 0,003943 \Delta \text{Gasoli}_{t-9} + 256,9 \Delta \text{Dias}_t + 437,0 S_{1t} \\ & - 611,1 M_{2t} + 909,9 M_{5t} - 532,7 M_{6t} + 193,1 M_{8t} + 531,8 M_{9t} \\ & - 364,0 M_{12t} + \varepsilon_t \end{aligned}$$

cu_j

os coeficiente de determinação e coeficiente de determinação ajustado são, respectivamente:

$$R^2 = 0,901 \text{ e}$$

$$R^2 \text{ ajustado} = 0,895.$$

Através dos habituais testes efectuados aos resíduos do modelo estimado constatámos a ausência de heterocedasticidade (*p-value* do teste de White é igual a 0,515) e a normalidade dos erros (*p-value* do teste de Jarque-Bera é igual a 0,068)¹⁹.

Quanto à autocorrelação, é significativa nos desfasamentos 4 a 6, a um nível de significância de 5% mas não o é se tomarmos aquele nível a 1%. Tentámos corrigir este problema através da inclusão de mais alguns desfasamentos das variáveis exógenas e endógena mas sem sucesso²⁰.

Apesar da introdução das variáveis explicativas, continuam a ter muita importância os desfasamentos da variável dependente, o número de dias do mês e as "dummies sazonais".

Os coeficientes estimados das variáveis indicadoras da conjuntura económica são significativos, o que mostra que essa conjuntura influencia o número de nascimentos que ocorrem cerca de nove meses depois. Os seus sinais estão de acordo com o que seria de esperar. Uma expansão da economia reflectida em variações positivas daqueles indicadores explicará, em parte, um aumento das intenções de procriar que, por sua vez, fará com que ocorram mais nascimentos nove meses depois.

A sazonalidade dos nados vivos é determinada pelo número de dias de cada mês e, também, por factores não identificados que, julgamos, se prendem com as preferências dos futuros pais quanto ao mês de nascimento das crianças e cujo efeito é considerado através das "dummies".

¹⁹ Os testes referidos encontram-se descritos em Greene (1993).

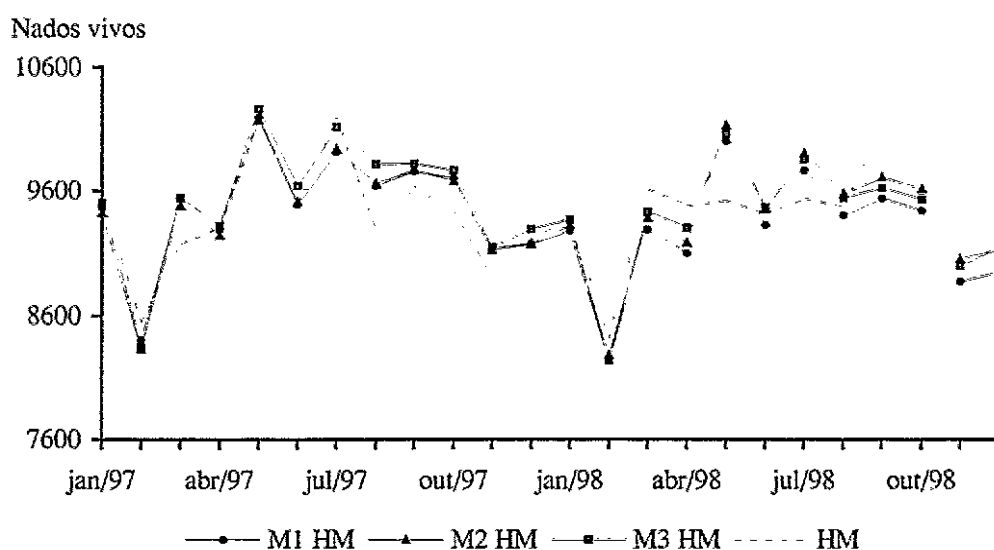
²⁰ Para testar a autocorrelação, utilizámos o teste de Breusch-Godfrey (Greene, 1993).

6. Comparação dos resultados dos vários modelos

Ao longo das três secções anteriores, experimentámos outros tantos métodos de previsão diferentes com vista a verificar se algum deles se destacava significativamente.

Comparemos, agora, os resultados obtidos, para os anos de 1997 e 1998.

As previsões são muito parecidas, todos os modelos produziram o mesmo padrão de sazonalidade.



É muito difícil distinguir entre as previsões dos modelos SARIMA e estrutural para o ano de 1997, sendo que as do modelo econométrico diferem apenas, e não muito significativamente, para os meses de Junho a Dezembro.

Quanto ao ano de 1998, a ideia com que ficamos pela análise do gráfico acima é que são os modelos estrutural e econométrico que quase se equivalem, não sendo, contudo, as previsões do SARIMA muito diferentes das fornecidas por aqueles.

Qualquer um dos modelos teria previsto, no fim de 1996, um aumento para o ano seguinte mais significativo do que o que aconteceu e, para o ano de 1998, a retoma da tendência decrescente (em 1997 nasceram vivas 112933 crianças, mais do que o previsto para o ano seguinte) que não terá acontecido realmente.

As previsões para 1997 do modelo SARIMA e do estrutural, como já dissemos, não diferem muito, sendo as deste ligeiramente melhores para o total e para o sexo masculino e piores para o sexo feminino.

O modelo econométrico, para aquele ano, destaca-se pela negativa em todos os aspectos: erro total e erro quadrático médio para os dois sexos e total.

Para o primeiro semestre de 1998, os dois primeiros modelos quase não diferem em termos de erro quadrático médio; porém, o estrutural cometeu um erro total menor (o que significa que os desvios mensais se anulam).

Tabela 15: Erros das previsões dos vários modelos para o ano de 1997 e 1º Semestre de 1998.

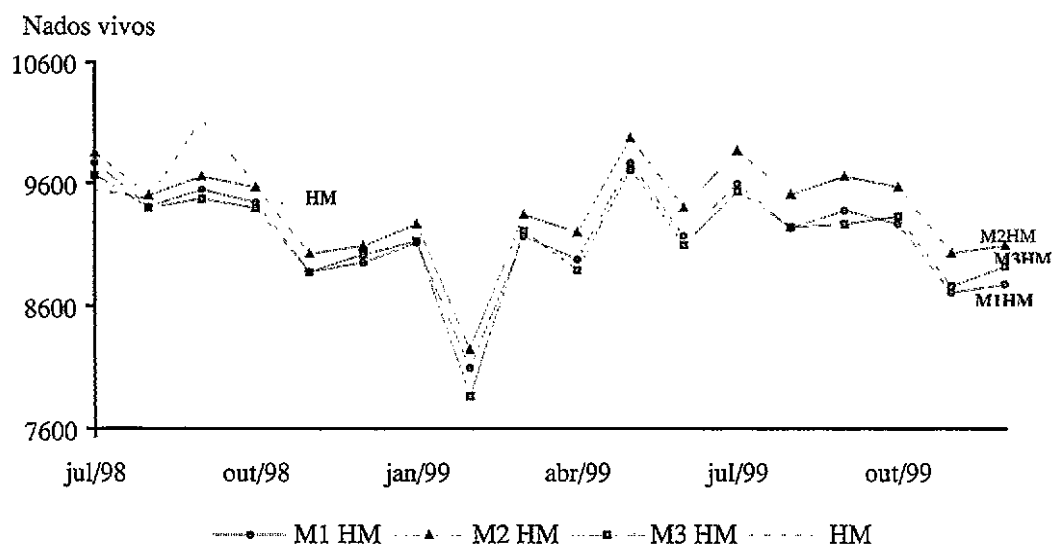
	Modelo1			Modelo 2			Modelo 3		
	H	M	HM	H	M	HM	H	M	HM
1997									
Total	58 656	55 065	113 721	58 533	54 951	113 484	59 110	55 491	114 601
Erro	-619	-169	-788	-496	-55	-551	-1 073	-595	-1 668
\sqrt{EQM}	121	110	205	114	111	198	140	125	243
1998									
Total	57 377	53 862	111 239	58 173	54 612	112 785	58 084	54 522	112 606
1º Sem.	28 498	26 752	55 250	28 777	27 014	55 791	28 823	27 056	55 879
Erro	441	175	616	162	-87	75	116	-129	-13
\sqrt{EQM}	170	143	300	163	151	301	138	133	256

O terceiro modelo destaca-se, aqui, pela positiva, dos anteriores.

O nosso objectivo principal é a previsão e o desejo, ao construir um modelo econométrico em que temos em conta, para além das observações passadas da variável em estudo, outras variáveis explicativas, era o de que este viesse a ser um melhor preditor do que os anteriores. Tendo em conta os resultados analisados acima, isso não é o caso do modelo a que chegámos. Contudo, o facto de, em relação a um período mais recente, o último modelo ter superado os outros deveria levar-nos a optar por escolhê-lo para efectuar as nossas previsões.

Comparemos, ainda, as previsões para o período de Jul/98 a Dez/99

Gráfico 36: Previsões segundo os vários modelos e dados provisórios para o período de Julho de 1998 a Dezembro de 1999



Os resultados fornecidos pelo último modelo quase igualam os alcançados através da metodologia ARIMA o que denota a indiferença na escolha de um deles.

Por seu turno, o modelo estrutural prevê que o nível dos nascimentos se situe mais acima, o que, de facto, terá acontecido.

7. Um método alternativo de obtenção de estimativas do número de nados vivos

Nas secções anteriores, descrevemos várias tentativas de encontrar um modelo que forneça boas previsões para o número de nados vivos de cada sexo ocorridos por mês.

A opção pela previsão desses valores referentes a meses passados quando ainda não existem dados disponíveis pareceu-nos não ter alternativa. Contudo, essa alternativa existe e passa, não por modelos de previsão algo complicados e, podemos dizê-lo agora, pouco fiáveis, mas por um método muito simples de obtenção de estimativas.

7.1. As estimativas do Instituto de Genética Médica para os nados vivos anuais

O Instituto de Genética Médica Jacinto de Magalhães (IGM) é responsável pelo rastreio efectuado aos recém-nascidos com vista ao diagnóstico precoce de doenças que podem ser evitadas se detectadas nos primeiros dias de vida.

Esse rastreio é feito através do chamado "teste do pezinho" até cerca de oito dias após o nascimento da criança. Não sendo obrigatório, tem-se verificado que a taxa de cobertura deste teste é bastante elevada: em 1996, foram rastreadas 108740 das 110243 crianças nascidas no nosso país, 98,637%; no ano seguinte, estudaram-se 111396 de um total de 112933, 98,639%.

Já há uns anos, no IGM, notou-se que a taxa de cobertura tinha vindo a aumentar, situando-se, então, em valores muito elevados. Chegou-se a um nível em aquela taxa varia pouco de ano para ano pois, por um lado, não se espera que decresça e, por outro, não pode aumentar muito mais por se situar perto dos 100%.

Esta constatação levou a que, naquele instituto, comesçassem a ser calculadas, no fim de cada ano, por estarem imediatamente disponíveis os números do rastreio, estimativas para o número de nascimentos com vida ocorridos nesse ano cujos valores oficiais, como sabemos, não assumem um carácter definitivo antes de meados do ano seguinte.

Sendo,

HM_n^E = número de crianças nascidas e estudadas no ano n e

HM_n = número de crianças nascidas no ano n ,

temos

$$TC_n = \frac{HM_n^E}{HM_n} \times 100, \text{ a taxa de cobertura do teste no ano } n.$$

A estimativa para o número de nascimentos ocorridos no ano n é:

$$\hat{HM}_n = \frac{HM_n^E}{\hat{TC}_n},$$

em que \hat{TC}_n é uma estimativa para a taxa de cobertura do teste nesse ano.

O IGM estima que a percentagem de recém-nascidos que são rastreados num ano seja a mesma do ano anterior, pelo que a sua estimativa para o número de nascimentos com vida é:

$$\hat{HM}_n = \frac{HM_n^E}{TC_{n-1}}.$$

Os dados referentes às crianças estudadas e às crianças nascidas nos últimos três anos, permitem-nos calcular as estimativas para 1997 e 1998.

Tabela 26: Crianças nascidas e estudadas nos últimos três anos.

n	HM_n^E	HM_n	TC_n	\hat{TC}_n	\hat{HM}_n
1996	108 740	110 243	98,637	-	-
1997	111 396	112 933	98,639	98,637	112 935
1998	112 210	-	-	98,639	113 758

Como a taxa de cobertura de 1997 se aproximou muito da do ano anterior, o número de dados vivos estimado para este ano quase igualou o verdadeiro valor.

Para o ano de 1998, estimou-se que terão nascido 113758 crianças, o que significa que continuou a tendência crescente iniciada em 1996.

7.2. As estimativas mensais e por sexo

Pretendemos, não só uma estimativa para o total de nascimentos ocorridos num ano mas a sua desagregação por meses e por sexos.

Sendo

$H_{n,i}$ = número de crianças do sexo masculino nascidas no mês i do ano n , $i = 1, \dots, 12$ e

$M_{n,i}$ = número de crianças do sexo feminino nascidas no mês i do ano n , $i = 1, \dots, 12$,

as suas estimativas serão,

$$\hat{H}_{n,i} = \frac{H_{n,i}^E}{\hat{TC}_{n,i}^H} \text{ e } \hat{M}_{n,i} = \frac{M_{n,i}^E}{\hat{TC}_{n,i}^M}$$

em que $H_{n,i}^E$ e $M_{n,i}^E$ são o número de crianças estudadas de cada sexo no mês i do ano n e $\hat{TC}_{n,i}^H$ e $\hat{TC}_{n,i}^M$ as estimativas das taxas de cobertura.

Tomaremos como estimativa para as taxas de cobertura mensais de cada sexo a taxa de cobertura total do ano anterior: $\hat{TC}_{n,i}^H = \hat{TC}_{n,i}^M = TC_{n-1}$.

Deste modo, é possível estimar com alguma precisão os valores pretendidos, porém, os erros cometidos serão mais significativos do que os da secção anterior pois as taxas de cobertura mensais, por sexos, variam mais do que as anuais.

Para calcular as estimativas acima apresentadas basta que existam os dados necessários, ou seja, a desagregação do número de crianças estudadas num ano por sexo e data de nascimento.

Solicitámos essa informação ao IGM que pôde disponibilizá-la uma vez que, para cada teste, é feito um registo do qual constam esses dados.

É significativo o número de fichas incompletas, às quais falta a data de nascimento ou o sexo, e não deve ser desprezado (em 1998, 2084 fichas não continham informação sobre o sexo da criança e 1085 não referiam a sua data de nascimento). Vamos admitir que existe a mesma probabilidade de que não seja referida a data de nascimento da criança seja qual for o mês em que nasceu (o que corresponde a distribuir o número de "datas desconhecidas" segundo a proporção de cada mês no total de "datas conhecidas"). Também consideraremos equiprovável a não referência ao sexo pelo que a cada um será atribuída uma percentagem de crianças desconhecidos igual à de sexo conhecido.

Introduziremos, aqui, outro tipo de erro que será tanto maior quanto mais as hipóteses admitidas se afastarem da realidade.

As estimativas pretendidas serão, então:

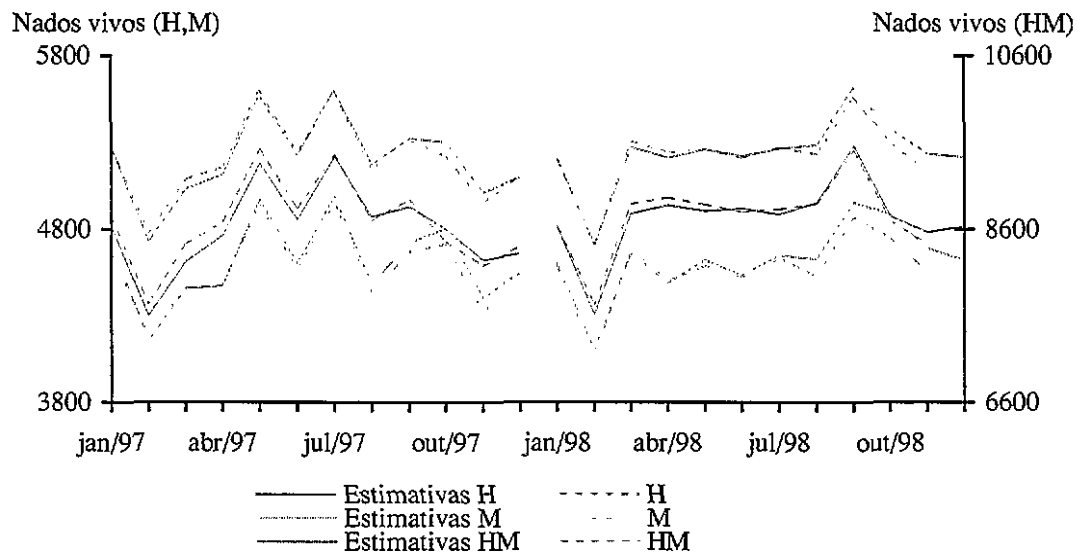
$$\hat{H}_{n,i} = \frac{\hat{H}_{n,i}^E}{TC_{n-1}} \text{ e } \hat{M}_{n,i} = \frac{\hat{M}_{n,i}^E}{TC_{n-1}}$$

em que $\hat{H}_{n,i}^E$ e $\hat{M}_{n,i}^E$ resultam das distribuições acima descritas.

Os dados que possuímos (para os últimos três anos) permitem-nos estimar o número de nascimentos com vida ocorridos por mês e por sexo nos anos de 1997 e 1998.

No gráfico seguinte encontram-se representadas as estimativas para os dois

Gráfico 27: Estimativas para os nados vivos para os anos de 1997 e 1998



anos, os dados para o ano de 1997 e dados provisórios para o ano de 1998.

As estimativas obtidas afastam-se pouco do número de nascimentos ocorridos para os meses do ano de 1997 e primeiro semestre de 1998. Os erros são mais significativos na estimação por sexos.

Tabela 18 : Erros das estimativas

	Estimativas			Dados		
	H	M	HM	H	M	HM
1997						
Total	57 617	55 318	112 935	58 037	54 896	112 933
Erro	420	-422	-2			
\sqrt{EQM}	62	48	80			
1998						
Total	28 761	26 999	55 760			
Até Nov.	53 528	50 791	104 319	53 609	50 253	103 862
1º Sem.	28 763	26 997	55 760	28 939	26 927	55 866
Erro	176	-70	106			
\sqrt{EQM}	40	26	41			

Como seria de esperar, os erros cometidos são quase insignificantes. As diferenças entre o número de crianças nascidas por mês, por sexo, são mais significativas do que o erro total anual devido ao maior erro da taxa de cobertura estimada e também ao facto de o número de crianças estudadas considerado ser uma estimativa.

As estimativas mensais para o ano de 1997 erraram, no máximo, cerca de 2%. No primeiro semestre de 98 nunca se atingiu esse erro percentual.

8. Conclusão

Ao longo das secções 3, 4, e 5, experimentámos três métodos de previsão diferentes tendo como objectivo encontrar um que fornecesse boas previsões.

Constatámos que as três metodologias podem fornecer modelos que funcionam como preditores razoáveis que nos teriam levado a cometer erros não muito significativos mas não esqueçamos que todos eles teriam previsto o retomar da tendência decrescente no ano passado. Na verdade, em 1998, aconteceu, pelo terceiro ano consecutivo, um aumento do número de nascimentos com vida no nosso país.

As melhorias que desejávamos alcançar com a modelização estrutural e com o modelo econométrico não aconteceram em grande medida (ver secção 6), verificámos que as previsões por eles fornecidas não diferem muito.

Há que ter em conta que atravessamos um período de instabilidade da natalidade, em que os modelos dificilmente poderiam ter um melhor desempenho.

É muito complicado, senão impossível, identificar os verdadeiros factores determinantes da evolução da natalidade e, mais ainda, quantificá-los e medir a sua influência.

O que se pretende fazer não é uma verdadeira previsão, no sentido de estimar valores relativos a algo que ainda não aconteceu, mas sim a previsão de valores referentes a períodos passados quando os dados oficiais ainda não estão disponíveis.

O que julgámos ser inevitável no início deste estudo, a opção pela previsão, acabou por se nos afigurar perfeitamente dispensável, quando surgiu a oportunidade de utilizar um método alternativo de cálculo de estimativas.

Esse método, ainda que bastante rudimentar, fornece resultados muito satisfatórios, pois tem por base números que são, eles próprios, uma boa aproximação dos que pretendemos estimar.

A opção não pode ser outra que não a de estimar os valores pretendidos segundo o método apresentado na secção 7.

Contudo, se pretendêssemos fazer verdadeiras previsões, poderíamos optar por um dos modelos, quase indiferentemente, sem que os resultados obtidos devessem merecer muita confiança, pelo menos enquanto não se verificar, novamente, estabilidade ao nível da tendência da natalidade.

Referências Bibliográficas

- Barata, O. S., *Natalidade e Política Social em Portugal*, I.S.C.S.P., Lisboa, 1985.
- Box, G. E. P., Jenkins, G. M. e Reinsel, G. C., *Time Series Analysis, Forecasting and Control*, 3ª Ed., Prentice Hall, Nova Jersey, 1994.
- Carrilho, M. J. , "Aspectos demográficos e sociais da população portuguesa no período 1864-1981: evolução global do continente português", *Estudos Demográficos*, nº 30, I.N.E., Lisboa, 1991.
- Carrilho, M. J. e Peixoto, J., "A evolução demográfica em Portugal entre 1981 e 1992", *Estudos Demográficos*, nº 31, I.N.E., Lisboa, 1993.
- Corrêa, A. C. M., "A Masculinidade nos nascimentos em Portugal", *Revista do Centro de Estudos Demográficos*, nº3, Lisboa, 1946.
- Costa, A. A., *Modelos Estruturais para Sucessões Cronológicas: uma apresentação*, Texto de Apoio nº7, CEMAPRE, 1993.
- Greene, W. H., *Econometric Analysis*, Prentice-Hall, New Jersey, 1993.
- Harvey, A. C. e Shepard, N., "Structural Time Series Models", G. S. Madalla, C.R. Rao e H. D. Vinod, eds., *Handbook of Statistics*, Vol.11, Elsevier Science Publishers, 1993.
- Harvey, A. C., "Trends Cycles and Autoregressions", *The Economic Journal*, nº107, Royal Economic Society, Oxford, 1997.
- I.N.E., *Anuários Demográficos, Estatísticas Demográficas e Séries de Estimativas Provisórias*.
- Koopman, S. J., Harvey, A. C., Doornik, J. A. e Shepard, N., *Stamp 5.0- Structural Time Series Analyser, Modeller and Predictor (Manual)*, Chapman & Hall, Londres, 1995.

Murteira, B., Müller, D. e Turkman, K. F., *Análise de Sucessões Cronológicas*, McGraw-Hill, Lisboa, 1993.

Nazareth, J. M., *Introdução à Demografia*, Editorial Presença, Lisboa, 1996.

Pressat, R., *Démographie Statistique*, Presses Universitaires de France, Paris, 1972.

Rosa, M. J. V., "O envelhecimento e as dinâmicas da população portuguesa a partir de 1960", *A Situação Social em Portugal, 1960-1995*, I.C.S., Lisboa, 1996.

Shryock, H. e Siegel, J., *The Methods and Materials of Demography*, Academic Press, Nova Iorque, 1976.

INFORMAÇÕES

VOLUME 3

3º QUADRIMESTRE DE 1998

ACTIVIDADES E PROJECTOS IMPORTANTES NO ÂMBITO DO SISTEMA ESTATÍSTICO NACIONAL

IMPORTANT ACTIVITIES AND PROJECTS IN THE SCOPE OF THE NATIONAL STATISTICAL SYSTEM

INQUÉRITO À OCUPAÇÃO DO TEMPO (IOT)

1. Introdução

Quando falamos de ocupação do tempo, estamos a referir-nos a que tempo? Ao tempo de trabalho, ao tempo de lazer, ao tempo que cuidamos da família, ou ainda ao tempo que passamos nas deslocações?

O tempo em que estamos interessados, são as 24 horas do dia, ou seja, o conjunto de actividades quotidianas das pessoas que vivem em Portugal, tratando-se principalmente de medir os tempos passados nas diversas actividades que compõem a vida de todos os dias: dormir, comer, trabalhar em casa, num escritório ou numa obra, jogar futebol, ir à pesca, ir ao café, ver televisão, tudo isto faz parte do universo de cada um de nós, sem que nos dêmos conta da parte que realmente cada uma delas ocupa na nossa vida.

O estudo do emprego do tempo, situa-se pois, no conjunto dos inquéritos sobre as condições e o modo de vida das famílias e dos indivíduos, estando o respectivo inquérito para os comportamentos não monetários, como o inquérito aos orçamentos familiares está para os comportamentos monetários.

2. Breve história do projecto

No campo das estatísticas sociais, os inquéritos sobre a ocupação do tempo, têm já tradição, sendo realizados com periodicidade regular em alguns países da União Europeia, bem como noutros, como é o caso do Japão e do Canadá (e da *ex*-União Soviética).

Todavia e tal como acontece em outras áreas estatísticas, a comparabilidade dos dados resultantes destes inquéritos é bastante limitada, em consequência de metodologias diversas utilizadas em cada um dos países.

No âmbito comunitário, e na sequência do respectivo Grupo de Trabalho dos Indicadores Sociais, foi decidido avançar com um estudo para avaliação da possibilidade de realizar um inquérito harmonizado, tendo-se para o efeito constituído uma Task-Force, sendo neste contexto que Portugal se insere, participando na Task-Force e realizando em 1996 um inquérito piloto, em conjunto com os demais países da União Europeia, tendo o Eurostat, enquanto coordenador do projecto, preparado a metodologia e os questionários.

Tendo em conta o trabalho técnico já desenvolvido e a importância do projecto, considerou-se que se devia avançar para a sua realização, tendo-se decidido ir ao exterior procurar parceiros que viabilizassem a sua concretização. A riqueza da informação que o inquérito proporciona, aliada à possibilidade de introduzir "módulos" enquadráveis no tema do projecto, conduziram a que algumas entidades públicas e do sector privado aceitassem o convite .

É assim que organizações ligadas ao sector do áudio-visual e dos *mass* média, bem como dos transportes, da cultura, da distribuição e da problemática da família, tivessem aderido ao projecto, financiando em diferentes proporções a sua execução. A par de entidades co-financiadoras, o INE pôde ainda contar com a colaboração técnica quer do Instituto de Ciências Sociais e do Observatório das Actividades Culturais, no que respeita particularmente à análise dos dados.

3. Objectivos

O Inquérito à Ocupação do Tempo tem como principal objectivo disponibilizar indicadores importantes para a escolha de políticas adequadas à realidade presente e que se perspectiva para futuro.

Assim, elegeram-se como objectivos específicos principais os seguintes:

- a) Caracterizar o padrão actual da ocupação do tempo da população portuguesa com seis ou mais anos (actividades, durações);
- b) Estudar o apoio que as famílias recebem, bem como o apoio que cada indivíduo oferece não só a outras famílias como também a organizações sem fins lucrativos, como trabalho voluntário (humanitárias, políticas, culturais e recreativas, etc.);
- c) Pesquisar as opiniões dos indivíduos sobre a sua percepção relativamente à organização do seu tempo;
- d) Fornecer dados sobre as deslocações quotidianas, durações e meios de transportes utilizados;
- e) Contribuir para a formulação de políticas ligadas à família, como sejam as relações entre o trabalho e as responsabilidades familiares, e a proporção crescente de mulheres na vida activa; os cuidados às crianças aos incapacitados e idosos.

4. Variáveis observadas e questionários

O tipo de informação a recolher é um compromisso entre vários vectores: por um lado o projecto harmonizado comunitário e, por outro, a necessidade de conseguir um equilíbrio entre os recursos financeiros e o que se considerou que deveria ser o fundamental da informação a obter.

Assim, relativamente ao modelo comunitário, foram feitas algumas simplificações, particularmente na caracterização dos activos empregados.

São utilizados três questionários, podendo a informação obtida através de cada um deles ser cruzada com a restante: Questionário de Família; Questionário Individual; Diário Individual.

Desde logo o **Diário** surge como a peça chave da recolha da informação, sendo neste questionário, que cada um dos inquiridos vai “dizer” o que fez, ou seja qual a **actividade principal** realizada, o que fez enquanto **actividade simultânea**, o **local** em que essa actividade principal se realizou e **com quem** estava.

Todo este descritivo é feito para intervalos de tempo de 10 minutos, começando às 24 horas, às 4 horas da manhã e terminando às 4 horas do dia seguinte.

Posteriormente, o texto que os inquiridos escreveram, é “traduzido em códigos”, de acordo com uma nomenclatura que foi criada para este efeito, e que está harmonizada ao nível comunitário, estando organizada em 10 classes, cada uma delas desagregada a três dígitos, cobrindo as seguintes actividades:

- cuidados pessoais;
- emprego;
- estudo;
- trabalhos domésticos e cuidados à família;
- actividades cívicas e de voluntariado;
- vida social e entretenimento;
- prática de desporto;
- passatempos e jogos;
- meios audio-visuais;
- trajectos.

CÓD. I. N. E.	HORAS	O QUE ESTAVA A FAZER? Indique a sua ocupação/actividade principal, em cada período de 10 minutos. (Indique apenas uma ocupação/actividade, em cada linha)	CÓD. I. N. E.	O QUE A FAZER PARA MÁIS ESTAVA ALÉM DO QUE JÁ INDICOU NA COLUNA 3? (Indique máis a ocupação importante)	CÓD. I. N. E.	ESTAVA SOZINHO OU COM ALGUÉM QUE CONHECE? (Indique com X)				LOCALIZADO DE TRANSPORTE Casa, residência, trabalho, loja, supermercado, banco, etc. (etc.)	CÓD. I. N. E.
						Solteiro	Com alguém até 14 anos que vivam no mesmo espaço	Com alguém residente no mesmo espaço que não seja colega de casa	Com alguém que esteja no mesmo espaço que não seja colega de casa		
1	2	3	4	5	6	7	8	9	10	11	12
<input type="checkbox"/>	4:00-4:10		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	4:10-4:20		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	4:20-4:30		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	4:30-4:40		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	4:40-4:50		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	4:50-5:00		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	5:00-5:10		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	5:10-5:20		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	5:20-5:30		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	5:30-5:40		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	5:40-5:50		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	5:50-6:00		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	6:00-6:10		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	6:10-6:20		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	6:20-6:30		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	6:30-6:40		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	6:40-6:50		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
<input type="checkbox"/>	6:50-7:00		<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

O **Questionário Individual** permite caracterizar os indivíduos em termos da sua relação com o trabalho, a par de outras questões sobre a avaliação do ritmo de vida e apoio a outras famílias ou a organizações, estando ainda incluídos alguns temas relacionados com a participação dos patrocinadores do projecto: “Televisão”, “Cinema”, “Leitura”, “Actividades socio-culturais”, “Partilha de actividades”.

O **Questionário de Família** faz o enquadramento dos indivíduos na família a que pertencem, ou seja, dimensão e tipologia, bem como algumas questões ligadas ao apoio que a família recebe em várias vertentes (de outras famílias, na compra de alguns serviços) e ainda na caracterização da utilização da televisão e do computador, sendo também neste questionário que se obtêm informação sobre as características demográficas dos indivíduos e sobre as fontes e montantes do rendimento familiar.

5. Síntese metodológica

5.1. Âmbito Geográfico

O IOT abrange o Continente e as Regiões Autónomas dos Açores e da Madeira, tendo-se estabelecido que os resultados devem apresentar representatividade ao nível de NUTS 1 e NUTS 2 e das Áreas Metropolitanas de Lisboa e do Porto para as principais variáveis previstas no inquérito.

5.2. População

A população em estudo é constituída pelos indivíduos, do sexo masculino ou feminino, residentes em unidades de alojamento não colectivas, com 6 ou mais anos de idade.

5.3. Unidade estatística

A unidade estatística principal a observar neste inquérito é o Indivíduo com idade acima do limiar já referido, enquanto que a família apenas funciona enquanto elemento caracterizador do contexto em que o(s) indivíduos se encontra(m) inserido(s).

Esta metodologia tem como consequência, que não são inquiridos todos os membros da família, mas apenas um ou dois, de acordo com as seguintes regras:

Tipo de família	Tipo e n.º de pessoas a inquirir
-Constituída por 1 pessoa	-Um adulto (M ou F)
-Constituída por 2 ou mais pessoas	-Um adulto e uma criança ou dois adultos

5.4. Amostra

A base utilizada para a amostragem foi a Amostra-Mãe do INE, que é uma amostra de unidades de alojamento, destinada a ser utilizada na selecção de amostras para inquéritos às famílias, tendo sido adoptado o método de amostragem probabilística e multietápica (3 etapas).

Na primeira etapa seleccionaram-se as freguesias, por constituírem as unidades administrativas mais pequenas, com limites físicos bem definidos.

Na segunda etapa, seleccionaram-se, nas freguesias da amostra, secções estatísticas (áreas geográficas contíguas, delimitadas cartograficamente, com cerca de 300 alojamentos).

Na terceira etapa, seleccionaram-se alojamentos até perfazer a dimensão da amostra previamente definida.

A dimensão da amostra foi fixada em cerca de 5500 unidades de alojamento, para um total de 10 000 indivíduos a inquirir, que incluem reforços para as áreas metropolitanas de Lisboa e Porto, de modo a garantir, também, representatividade para estas áreas.

6. Recolha da informação

De acordo com o calendário estabelecido, o trabalho de campo da recolha directa da informação decorre nos meses de Outubro e Novembro de 1999, sendo realizada por 145 agentes entrevistadores, espalhados por todo o território nacional, sendo enquadrados por uma equipa de supervisores.

Tendo em conta as características do inquérito, a recolha da informação de base é efectuada em dois momentos:

1º momento: entrevista à família/indivíduo, efectuando a recolha dos questionários de família e individual e explicando o preenchimento do diário;

2º momento: recolha dos diário e/ou apoio ao respectivo preenchimento, quando necessário.

7. Tratamento da informação e divulgação

A informação obtida sofre internamente um processo de verificação/codificação, registo e validação correndo em paralelo o controle de qualidade.

Também no que respeita à recolha, é realizado um controle de qualidade, quer sobre a qualidade da informação obtida, quer sobre a utilização da metodologia proposta.

O manancial de informação que se irá obter através deste inquérito, permite equacionar uma divulgação que se não esgota numa única publicação mas dando preferência a diversas publicações.

A primeira divulgação será efectuada no final do 1º semestre de 2000, e da qual constarão dados de âmbito geral, admitindo-se a execução de várias outras publicações que serão mais detalhadas cobrindo aspectos mais específicos do inquérito.

CONGRESSOS, SEMINÁRIOS, COLÓQUIOS E CONFERÊNCIAS

CONGRESS, SEMINARS AND CONFERENCES

2000

□ 17-19 de Janeiro

Annual Meeting on "Statistics and Planning to Help Young People in Big Agglomerations" ("Statistik und Jugendhilfeplanung in der Großstadt"), Berlin, Germany (Haus am Rupenhorn).

The aim of this smaller international conference of the Federal Statistical Office, the Statistisches Landesamt Berlin and the Haus am Rupenhorn is to bring together scientists from research institutes, official statisticians and social workers from different countries in a former private home of a Jewish family above the Havel river. There we try to learn from each other about the problems of the younger generation and we are looking together for possible solutions.

Informações: Prof. Dr. Eckart Elsner, Statistisches Landesamt Berlin,
Alt-Friedrichsfelde 60, 10306 Berlin, Germany.
Tel.: +49-30-9021-3488.
Fax: 49-30-9021-2561.
E-mail: e.elsner@statistik-berlin.de

□ March

7th Annual Meeting of the Portuguese Association of Classification and Data Analysis (CLAD), Lisbon, Portugal.

Official languages are Portuguese, French and English, as usual.

Informações: E-mail clad.lead@mail.telepact.pt or cladlead@fc.ul.pt

□ 8-9 March

9th Dedicated ASIM Conference "Experiences from the Future — Simulation in Production and Logistics", Berlin, Germany.

Organiser: Fraunhofer Institute for Production Systems and Design Technology, Pascalstrasse 8-9, D-10587 Berlin.

Informações: Markus Rabe. Tel.: +49-30-3900-6248. Fax: +49-30-393-2503.
URL: <http://www-plt.ipk.fhg.de/ASIM-Fachtagung>
E-mail: Markus.Rabe@ipk.fhg.de.

□ 16-19 March

Statistics: Reflections on the past and visions of the future. An international conference in honour of Professor C.R. Rao on the occasion of his 80th birthday University of Texas, San Antonio.

Informações: Associate Professor Nandini Kannan;
E-mail: kannan@sphere.math.utsa.edu

□ 17–18 March

The second bi-annual Western Statistics Teachers Conference (WeSTCo 2000: Statistics Education for the New Century) Greeley, Colorado, USA.

Informações: Dan Mundfrom: Tel.: +1-970-351-1669;

E-mail: djmund@unco.edu

URL: <http://www.math.armstrong.edu/faculty/lesser/WeSTCo2000.html>

□ 21–24 March

The fourth "Stochastik Tage" (German Open Conference on Probability and Statistics), Hamburg, Germany.

Informações: G. Neuhaus, Institut für Mathematische Stochastik, Universität Hamburg, 20146 Hamburg, Germany.

E-mail: stoch2000@math.uni-hamburg.de

URL: <http://stoch2000.math.uni-hamburg.de>

□ 13-15 April

Population Association of America Annual Meeting, Los Angeles, California, Westin Bonaventure.

Informações: PAA at 721 Ellsworth Dr., Suite 303, Silver Spring, MD 20910: (301) 565-6710 Fax: (301) 565-7850.

□ 13–16 April

Sunbelt XX International Sunbelt Social Network Conference, Costa Plaza Suite Hotel at Stanley Park, Vancouver, British Columbia, Canada.

Informações: URL: <http://www.sfu.ca/~insna>

□ 16–19 April

FRACTAL 2000. "Complexity and Fractals in the Sciences", 6th International Multidisciplinary Conference, Singapore.

Informações: URL: <http://www.kingston.ac.uk/fractal/>

□ 8-9 May

Short course entitled 'Interim Analyses in Clinical Trials', given by KyungMann Kim (Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison). Location : Catholic University Leuven, Belgium.

Informações: Bart Spiessens, Biostatistical Centre, Kapucijnenvoer 35,

B-3000 Leuven, Belgium

Tel: 32 / 16 / 33 68 96

Fax: 32 / 16 / 33 70 15.

E-mail: Bart.Spiessens@med.kuleuven.ac.be

URL: <http://www.med.kuleuven.ac.be/biostat/>

□ 14–16 May

Workshop on Hierarchical Modeling in Environmental Statistics, Ohio State University, Columbus, Ohio, USA.

Informações: Noel Cressie, Department of Statistics, 1958 Neil Avenue, 404 Cockins Hall, The Ohio State University,

Columbus, Ohio,

OH 43210-1247, USA.

Tel.: +1-614-292-5194;

Fax: +1-614-292-2096;

E-mail: ncressie@stat.ohio-state.edu

URL: <http://www.stat.ohio-state.edu/~ncressie>

□ 15–20 May

Fifth World Congress of the Bernoulli Society for Mathematical Statistics and Probability and 63rd Annual Meeting of the Institute of Mathematical Statistics, Guanajuato, Mexico.

Informações: IMS/Bernoulli Society, CIMAT,
Apdo. Postal 402, Guanajuato, GTO,
36000 Mexico.
Tel.: +52-473-27155;
Fax: +52-473-24511;
E-mail: wc2000@fractal.cimat.mx
URL: <http://www.Bernoulli-ims-2000.org.mx>

□ 15 - 19 May

Les 32 Journées de Statistique de la Société française de Statistique (SFdS), Maroc.

Informações: E-mail : sfds2000@fesnet.net.ma
URL: <http://www.fesnet.net.ma/sfds2000>

□ 17-19 May

CASI 2000 - Conference on Applied Statistics in Ireland - Twenty Years a'Growing - Fiche Blian ag Fas. The 20th Conference on Applied Statistics in Ireland, Great Southern Hotel, Rosslare, Co Wexford, Ireland under the auspices of the Irish Statistics Association.

Informações: Prof John Haslett, Dept of Statistics,
Trinity College, Dublin Ireland.
E-mail: John.Haslett@tcd.ie
URL: <http://www.tcd.ie/Statistics/CASI2000/CASI2000.htm>

□ 28 May–1 June

The 6th World Meeting of the International Society for Bayesian Analysis, Knossos Royal Village Hotel, Hersonissos, Crete, Greece.

Informações: Local organiser: Professor George Kokolakis,
National Technical University of Athens, Zografou Campus,
15780 Athens, Greece.
Tel.: +30-1-772-1704;
Fax: +30-1-772-1705;
E-mail: kokolakis@math.ntua.gr
Chair of Committee: Professor Mike West, Duke University, Durham,
NC 27708-0251, USA.
Tel./Fax: +1-919-684-8842/8594; E-mail: mw@stat.Duke.EDU
URL: <http://www.bayesian.org/isba2000/isba2000.html>
<http://www.ntua.gr/ISBA2000>

□ 4-7 June

Annual Meeting of the Statistical Society of Canada, Ottawa, Ontario, Canada.

Informações: Local Arrangements Chair: Andre Dabrowski, Department of Mathematics and Statistics,
University of Ottawa, 585 King Edward, Station A,
P.O. Box 450, Ottawa,
Ontario, K1N 6N5,
Canada.

E-mail: ardsg@mathstat.uottawa.ca

Program Chair: Duncan Murdoch, Department of Statistics and Applied Science, University of Western Ontario, London,
Ont., N6A 5B7,
Canada.

E-mail: murdoch@fisher.stats.uwo.ca

□ 5-9 June

International Conference on Advances in Convex Analysis & Global Optimization, Pythagorion, Samos, Greece.

Informações: Mrs. Thea Vigli-Papadaki, Department of Mathematics,
University of the Aegean,
Karlovassi 83200,
Samos, Greece.

Phone ++30-273-82021, 82010,

Fax ++30-273-82009

Email: acago@math.aegean.gr

URL: <http://www.samos.aegean.gr/math/acago>

□ 13-16 June

22nd International Conference on "Information Technology Interfaces" (ITI 2000), Pula, Croatia.

Informações: Tel.: +385-1-616-5597;

Fax: +385-1-616-5591;

E-mail: iti@srce.hr

URL: <http://www.srce.hr/iti>

□ 14-18 June

CHALLENGES FOR SCIENCE AND ENGINEERING in the 21st century, Stockholm, Sweden.

The conference is organized by the International Network of Engineers and Scientists for Global Responsibility, (INES), in cooperation with major organizations in the fields of science and engineering.

Informações: INES 2000 Conference Secretariat,
Gutenbergstr. 31, 44139 Dortmund, Germany.

Tel: +49 231 575218,

Fax: +49 231 575210,

E-mail: INES2000@t-online.de

URL: <http://www.ines2000.org>.

□ 17-21 June

ICES-II, International Conference on Establishment Surveys-II, Survey Methods for Businesses, Farms and Institutions, Adams Mark Hotel in Buffalo, New York.

Informações: URL: <http://www.eia.doe.gov/ices2/index.html>.

Alternatively, contact John G. Kovar at kovar@statcan.ca

□ 19-23 June

Fifth International Forum on Tourism Statistics, Glasgow, Caledonian University, Glasgow, UK.

Informações: Claire Scott, Moffat Centre for Travel & Tourism Business Development,
Glasgow Caledonian University,
1 Park Drive, Glasgow, G3 6LP,
U.K.

Tel.: +44-141-337-4224;

Fax: +44-141-3374222;

E-mail: c.a.scott@gcal.ac.uk

URL: <http://www.moffatcentre.com>

<http://www.tourism-conference.com>

□ 19-23 June

The Third International Workshop on Automatic Differentiation (AD), "AD2000 From Simulation to Optimization", Nice, Côte d'Azur, France.

Informações: URL: <http://www-sop.inria.fr/tropics/ad2000/>

□ 21 - 24 Junho

20th Internacional Symposium on Forecasting (ISF'2000), Lisboa

Informações: URL: <http://www.isf2000.org>

□ 21 - 24 juin

Société Française de Statistique - Colloque francophone sur les sondages-Bruxelles/Lille.

Informações: Jean-Jacques DROESBEKE, Université Libre de Bruxelles, LMTD -

CP 124, Avenue Jeanne 44,

B-1000 Bruxelles.

Tél. : (32-2) 650.32.74 ;

Fax : (32-2) 650.34.66.

URL: <http://www.ulb.ac.be/soco/lmtld/sondage2000/index.html>

E- mail : psemerar@ulb.ac.be

□ 25-30 juin

En liaison avec le colloque de Bruxelles-Lille, un cours de sondage pour les statisticiens francophones des pays en développement est prévu par l'Association Internationale des Statisticiens d'Enquêtes (AISE) et le CEFIL (INSEE, Libourne, France).

Informações: E-mail : christophe.lefranc@insee.fr

□ 29 June-1 July

International Conference on Statistics in the 21st century, University of Maine, Orono, ME, USA.

The conference will feature several prominent speakers. Professor C.R. Rao has agreed to be the key note speaker.

Informações: Professor Ramesh Gupta,

Department of Mathematics & Statistics, University of Maine, Orono,
ME, 04469-5752,
U.S.A.

Phone: 207-581-3913

Fax: 207-581-3902.

E-mail: RCGupta@maine.maine.edu

□ 2-7 July

The XXth International Biometric Conference, Campus of the University of California at Berkeley, USA.

Informações: URL: <http://www.biostat.ucsf.edu/IBC2000/>

□ 3-7 July

15th Australian Statistical conference, Adelaide, Australia.

Informações: R. Jarrett, Conference Convenor.

Fax: +61-8-8274-6000;

URL: <http://www.sapmea.asn.au/15ASC.htm>

□ 4-7 July

Second International Conference on Mathematical Methods in Reliability-MMR '2000, Bordeaux, France.

Informações: Dr. Valentina Nikoulina, Université Victor Segalen-Bordeaux 2,

UFR MI2S, B.P. 69,
33076 Bordeaux Cédex, France.

Tel.: +33-557-571-070;

Fax: +33-557-571-532;

URL: <http://www.mass.u-bordeaux2.fr/MI2S/MMR2000>

□ 11-14 July

IFCS 2000, 7th Conference of the International Federation of Classification Societies, Namur, Belgium.

Informações: Professor A. Hardy, Secretary Local Organizing Committee, Department of Mathematics,

The University of Namur, Rempart de la Vierge 8,
5000 Namur, Belgium.

E-mail: Andre.Hardy@fundp.ac.be

□ 17-19 July

The 22nd Biennial Conference of the Society for Multivariate Analysis in the Behavioural Sciences, Bankside House, 24 Sumner Street, London SE1 9JA, UK.

Informações: Dr Irini Moustaki.

Tel.: +44-207-955-6063;

Fax: +44-207-955-7416;

E-mail: smabs2k@lse.ac.uk

URL: <http://www.lse.ac.uk/statistics/smabs2k>

□ 17–21 July

15th International Workshop on Statistical Modelling. New Trends in Statistical Modelling, Bilbao, Spain.

Informações: Vicente Nunez–Anton, Eva Ferreira,

Departamento de Econometria y Estadística, Universidad del País Vasco, Avenida Lehendakari Aguirre 83,
48015 Bilbao, Spain.

Tel.: +34–94–601–3749/3739/3740;

Fax: +34–94–601–3754;

E-mail: vn@alcib.bs.ehu.es or: etpnuanv@bs.ehu.es or:

E-mail: ef@alcib.bs.edu.es or: etpfegae@bs.ehu.es

URL: <http://iwsn.bs.ehu.es>

□ 19–22 July

International Conference on Spatial Statistics in the Agro, Bio and Geosciences
Freiberg, Germany.

Informações: Dr Martina Lorenz, Freiberg

University of Mining and Technology, Graduate College on Spatial
Statistics, B.-v.-Cotta-Str. 2,
D-09599 Freiberg.

E-mail: lorenz@mailtuba.tu-freiberg.de

URL: <http://fink.mathe.tu-freiberg.de/conf.htm>

□ 24–28 Juillet

Un atelier sur les enquêtes en continu sur la force de travail aura lieu à Libourne au CEFIL à l'initiative de l'AISE et de l'INSEE du 24 au 28 juillet 2000. Nombre de places limité à 70. Traduction simultanée français-anglais envisagée.

E-mail : riandey@ined.fr

□ 1–5 August

ISI Cutting–Edge Conference on the Social and Biological Determinants of Longevity, Voorburg, The Netherlands, jointly organized by ISI and the Max Planck Institute for Demographic Research.

Informações: Professor James Vaupel, Director Max Planck Institute for Demographic Research,

Doberaner Strasse 114, 18057 Rostock, Germany.

E-mail: JWV@demogr.mpg.de

□ 7–11 August

IASE Round Table Conference on the Training of Researchers in the Use of Statistics, Institute of Statistical Mathematics, Tokyo, Japan.

Informações: Carmen Batanero, Departamento Didáctica de la Matemática, Facultad de Educación, Campus de Cartuja,
18071 Granada, Spain.

E-mail: batanero@goliat.ugr.es

URL: <http://www.ugr.es/~batanero/iasert.htm>

□ 13-17 August

Joint Statistical Meetings, Indianapolis, Indiana. Indiana Convention Center & RCA Dome. Sponsored by ASA, ENAR, WNAR, IMS, and SCC.

Informações: ASA, 1429 Duke St., Alexandria ,
VA 22314-3415;
Tel. (703) 684-1221;
Email meetings@amstat.org

□ 20-25 August

16th IFIP World Computer Congress, Beijing, China.

Informações: IFIP Secretariat, Hofstrasse 3,
2361 Laxenburg, Austria.
Tel.: +43-2236-73616;
Fax: +43-2236-736169;
URL: <http://www.ifip.or.at/> Ftp: <ftp.ifip.or.at>

□ 21-25 August

Compstat 2000, organized jointly by Statistics Netherlands and Utrecht University, Utrecht, The Netherlands.

Informações: Compstat 2000, FBU Congress Bureau, Utrecht University,
Attn: Mrs. Marcelle Buma,
P.O. Box 80125, 3508 TC Utrecht, The Netherlands.
Phone: +31-30-2532728
Fax: +31-30-2535851;
E-mail: compstat@fbu.uu.nl
URL: <http://neon.vb.cbs.nl/rsm/compstat>

□ 28-29 August

Lexis in Context: German and Eastern & Northern European Contributions to Demography 1860-1910.

Informações: Dr. Jochen Fleischhacker, Working Group on the History of Demographic Thinking, Max Planck Institute for Demographic Research,
Doberanerstr Str. 114, D-18057 Rostock.
Phone: (0049) 0381 2081-114.
Fax: (0049) 0381 2081-269.
Email: Fleischhacker@demogr.mpg.de

□ 4-8 September

21st ISCB (International Society for Clinical Biostatistics) Annual Meeting, Trento, Italy.

Informações: Fax +39-0461-231044
Email: ICSB2000@gelso.unitn.it
URL <http://www.gelso.unitn.it/~iscb2000>

□ 4–8 September

IAOS 2000 Conference on Official Statistics and Human Rights, Neuchâtel, Switzerland.

Informações: Ms. Carol Mottet, Department of International Affairs,
Swiss Federal Statistical Office, 10 Espace de l' Europe,
2010 Neuchâtel, Switzerland.
Tel.: +41–32–713–6009;
Fax: +41–32–713–6002;
E-mail: carol.mottet@bfs.admin.ch
URL: <http://www.admin.ch/bfs/about/international/iaos-htm>

□ 4–8 September

TIES/SPRUCE 2000, International Conference on Current Environmental Issues:
Quantitative Methods, University of Sheffield, United Kingdom.

Informações: URL: <http://www.shef.ac.uk/ties-spruce2000/>

□ 11-15 September

Royal Statistical Society International Conference, University of Reading, UK.

Informações: Dr Dave Collett,
ph +44 118 931 8024
fx +44 118 975 3169
E-mail d.collett@rdg.ac.uk
URL: <http://www.rdg.ac.uk/RSS2000>

□ 17–21 September

Open Conference on Spatial and Computational Statistics, Charlotte Mason
College, Ambleside, United Kingdom.

Informações: Cathy Thomson
Tel. +44-1524-594622
Email: c.thomson@lancaster.ac.uk
URL: <http://www.maths.lancs.ac.uk/~schlathe/ambleside.html>
(further details and on-line registration).

□ 18-27 September

**8th Silivri–G.Magusa (Famagusta) Workshop in Stochastic Analysis and related
topics**, Eastern Mediterranean University, G.Magusa, North Cyprus. Seminal lectures
will be delivered by P. Malliavin, D.W. Strook and A.Bensoussan (tentative).

Informações: Professor Ulug Capar;
E-mail capar.as@mozart.emu.edu.tr
URL: <http://mozart.emu.edu.tr/workshop>
<http://www.emu.edu.tr>

□ 18–22 September

International Data Analysis Conference (IDA 2000), Innsbruck, Austria.

Informações: Professor Reinhard Viertl, Technische Universität Wien, Institut für
Statistik,
Wiedner Hauptstr. 8/107, 1040 Wien, Austria.
Tel.: +43–1–58801;
URL: <http://www.statistik.tuwien.ac.at/ida2000>

- 20-22 September
Meeting on "The right of the citizen to be provided with information", Berlin, Germany.
Europäische Akademie,
Berlin, Germany.
- 25-29 September
German Statistical Week, Neurenberg, Germany.
Informações: Hans Teschner, Stadt Bielefeld, Betrieb Stadtforschung, Statistik und Wahlen, Niederwall 23,
33602 Bielefeld.
Tel.: +49-521-512108;
Fax: +49-521-513445;
E-mail: hteschn@aol.com
- 3-6 October
Fifth International Conference on Social Science Methodology of the Research Committee on Logic and Methodology (RC33) of the International Sociological Association (ISA), Cologne, Germany.
Informações: Joerg Blasius - Zentralarchiv für Empirische Sozialforschung
Universitaet zu Koeln - Bachemer Str. 40 - 50931 Koeln
Email: rc33@za.uni-koeln.de
URL: <http://www.za.uni-koeln.de/rc33>
- 28 October- 1 November
The Seventh Japan-China Symposium on Statistics
Informações: <http://www.e.u-tokyo.ac.jp/jcstat7>
- 7-10 November
22nd Conference on Regional and Urban Statistics and Research, Shenzhen, China.
Informações: Secretariat Office of the Organizing Committee, Shenzhen Statistics & Information Bureau, Add: No.1 Tongxin Rd.,
Shenzhen, P.R. China.
Tel. (86) 755-2240317
Fax (86) 755-2241559.
Email: yangwei@shenzhen.gov.cn
URL: <http://www.scorus2000.shenzhen.gov.cn>
- 9-13 December
The Ninth International Workshop on Matrices and Statistics in honour of Professor C R Rao's 80th birthday, in Hyderabad, India.
Regarding the scientific programme
Informações: contact P. Bhimasankaram, pbhim@hotmail.com S. Spuntanen, sjp@uta.fi K. Viswanath, kvsm@uohyd.ernet.in or H.J. Werner, werner@united.econ.uni-bonn.de.
Regarding registration, deadline for submitting papers and local accomodation
Informações: contact P. Bhimasankaram, pbhim@hotmail.com K. Viswanath, kvsm@uohyd.ernet.in B. Sidharth, birlasc@hd1.vsnl.net.in or R.J.R. Swamy, nhasan@ouastr.ernet.in

□ 29–31 December

International Conference Celebrating Prof. C.R. Rao's 80th Birthday, organised by The Indian Statistical Institute, to be held in Calcutta, India.

Informações: E-mail: cr80@isical.ac.in

□ 30 December–3 January 2001

International Conference on Recent Developments in Statistics and Probability and Their Applications, India International Centre, New Delhi, India.

This International Indian Statistical Association (IISA) Biannual Conference is organized by IISA India Chapter in collaboration with other Indian organizations.

Informações: Kanwar Sen, Department of Statistics,
University of Delhi, Delhi 110007, India.
Tel.: +91-11-723-1427 (home);
+91-11-725-6617 (office);
E-mail: dustats@del3.vsnl.net.in
URL: <http://www.stat.ohio-state.edu/~hnn/IISA.html>

2001

□ 10–14 June

Annual Meeting of the Statistical Society of Canada, Burnaby, British Columbia, Canada.

Informações: Local Arrangements Chair: Tim Schwartz,
Department of Mathematics and Statistics,
Simon Fraser University, Burnaby,
British Columbia, V5A 1S6, Canada.
E-mail: tim@cs.sfu.ca

□ 5–9 August

Joint Statistical Meetings, Atlanta, Georgia. Atlanta Marriott Marquis and the Atlanta Hilton & towers. Sponsored by ASA, ENAR, WNAR, IMS and SSC.

Informações: ASA, 1429 Duke Street, Alexandria,
VA 22314-3415;
(703) 684-1221;
Email meetings@amstat.org

□ 6–10 August

International Conference on "Extremes in Theory and Practice", Leuven, Belgium.

Informações: Jan Beirlant, University Center of Statistics, Katholieke Universiteit Leuven, De Croylaan 52B, 3001 Heverlee, Belgium.
Tel.: +32-16-322789;
Fax: +32-16-322831;
E-mail: jan.beirlant@wis.kuleuven.ac.be

□ 13–19 August

23rd European Meeting of Statisticians, Funchal, Island of Madeira, Portugal.

Informações: E-mail: Dinis.Pestana@fc.ul.pt

□ 22–29 August

International Statistical Institute, 53rd Biennial Session (includes meetings of the Bernoulli Society, The International Association for Statistical Computing, The International Association of Survey Statisticians, The International Association for Official Statistics and The International Association for Statistical Education), Seoul, Korea.

Informações: ISI Permanent Office, Prinses Beatrixlaan 428,
P.O. Box 950, 2270 AZ Voorburg, The Netherlands.
Tel.: +31-70-337-5737;
Fax: +31-70-386-0025;
E-mail: isi@cbs.nl or visit the Session website at
<http://www.nso.go.kr/isi2001>

2002

□ 23-29 June

The 8th International Vilnius Conference on Probability Theory and Mathematical Statistics, Vilnius, Lithuania.

Chair of Committee: Professor Vytautas Statulevicius, Institute of Mathematics and Informatics, Akademijos str. 4, 2600 Vilnius, Lithuania.

Informações: Dr. Aleksandras Plikusas, Institute of Mathematics and Informatics,
Akademijos str. 4, 2600 Vilnius, Lithuania.
Tel.: 370-2-729209;
Fax: 370-2-729209;
E-mail: conf@ktl.mii.lt

□ 11-15 August

Joint Statistical Meetings, New York, Hilton and Sheraton New York. Sponsored by ASA, ENAR, WNAR, IMS, and SCC.

Informações: ASA, 1429 Duke St., Alexandria, VA 22314-3415;
Tel. (703) 684-1221;
Email meetings@amstat.org

□ 19–23 August

24th European Meeting of Statisticians, Prague, Czech Republic.

2003

□ 10–20 August

International Statistical Institute, 54th Biennial Session (includes meetings of the Bernoulli Society, The International Association for Statistical Computing, The International Association of Survey Statisticians, The International Association for Official Statistics and The International Association for Statistical Education), to be held in Berlin, Germany.

Informações: ISI Permanent Office, Prinses Beatrixlaan 428,
P.O. Box 950, 2270 AZ Voorburg, The Netherlands.
Tel.: +31-70-337-5737;
Fax: +31-70-386-0025;

E-mail: isi@cbs.nl or visit the Session website at <http://www.isi-2003.de>

ACÇÕES DESENVOLVIDAS PELO INE NO ÂMBITO DA COOPERAÇÃO BILATERAL E MULTILATERAL

ACTIONS ACHIEVED BY NSI IN THE SCOPE OF BILATERAL AND MULTILATERAL COOPERATION

(DE 1 DE SETEMBRO A 31 DE DEZEMBRO DE 1999)

a) *Cooperação desenvolvida com os PALOP e Macau:*

Realizou-se no Instituto Nacional de Estatística (INE), em Lisboa, a X Reunião dos Directores-Gerais de Estatística dos PALOP, Portugal e Macau. Esta reunião, organizada pelo INE de Portugal e apoiada pela Cooperação Portuguesa, contou com a presença dos responsáveis máximos pelos INE de cada um dos países e do território de Macau, nomeadamente: Dr. Joaquim Flávio Couto de Angola, Eng. Francisco Tavares de Cabo Verde, Dr. Serifo Monteiro da Guiné- Bissau, Dra. Fátima Choi de Macau, Dr. João Dias Loureiro de Moçambique, Eng. Mário Coelho de São Tomé e Príncipe, sendo presidida pelo Eng. Carlos Corrêa Gago, Presidente do INE de Portugal. A reunião contou ainda com a participação de representantes do Instituto da Cooperação Portuguesa e do Gabinete de Relações Externas e Assuntos Comunitários do Ministério do Equipamento, Planeamento e Administração do Território. No decurso dos trabalhos, foram avaliados os programas de cooperação estatística em curso com estes países, bem como actualizadas as perspectivas estratégicas de desenvolvimento da cooperação estatística.

Na sequência desta reunião, e aproveitando a estada em Lisboa de todos os responsáveis máximos dos INE dos PALOP, realizou-se um seminário sobre "Difusão de Informação Estatística Oficial". Os trabalhos deste seminário incidiram sobre os seguintes temas : A Função da Difusão no elenco das Atribuições dos INE, As Relações entre os INE e os Órgãos de Comunicação Social e a análise da Situação e Perspectivas da Difusão nos INE dos PALOP, Portugal e Macau.

No quadro do projecto comum sobre Classificações, Conceitos e Nomenclaturas, realizou-se, no período em apreço, um estágio de desenvolvimento dos trabalhos da Classificação Nacional de Bens e Serviços (CNBS), tendo participado o técnico do INE de Angola, Dr. Manuel Caculo.

No âmbito do projecto-piloto para a implementação do Sistema de Contas Nacionais das Nações Unidas (SCN-93), foi realizada uma missão de assistência técnica, junto do INE de Moçambique, pela Dra. Teresa Ferreira, com o objectivo de desenvolver os materiais de apoio a acções de formação.

No período em apreço, teve lugar em Maputo o IIº Seminário Regional para os PALOP sobre Nomenclaturas e Classificações Económicas. Esta acção, realizada com o apoio financeiro do Programa COMSTAT da Comissão Europeia e de Portugal, nomeadamente através da participação do principal orientador dos trabalhos, Dr. Saraiva Aguiar , permitiu o confronto de experiências sobre a concepção e organização das Classificações Económicas dos PALOP, bem como das perspectivas de desenvolvimento em cada um dos países.

No período em apreço, e no âmbito da **Cooperação Bilateral com os PALOP**, foram realizadas as seguintes acções:

Angola

Na execução do programa de cooperação foram realizados, um estágio no âmbito do projecto Planeamento da Actividade Estatística Global e uma visita de trabalho sobre a organização e funcionamento do INE e do SEN (Sistema Estatístico Nacional).

Cabo Verde

Na execução do programa de cooperação foi realizada uma acção de assistência técnica ao Recenseamento Empresarial que incluiu o fornecimento de equipamento e software de actualização para sistema de rede do FUE (Ficheiro de Unidades Estatísticas) e uma missão para instalação das novas soluções informáticas. No decurso da missão, realizada pelo Dr. Fernando Carvalho, foi possível instalar uma rede informática interna no INE de Cabo Verde.

Em continuidade do projecto de Apoio Institucional ao INE de Cabo Verde foi realizada uma missão de assistência técnica, pelo Dr. A. Ferreira da Cunha, tendo sido efectuado o segundo estudo de Avaliação das Necessidades de Informação Estatística Oficial, e formuladas propostas sobre os parâmetros e procedimentos internos com vista à implementação do articulado sobre a transgressão estatística e sobre a regulamentação da recolha directa de informação estatística.

Teve lugar Lisboa, a 5ª reunião da Comissão Coordenadora da Gestão do Acordo de Cooperação Estatística Luso - Caboverdiano, na qual se estabeleceu a actualização do programa de cooperação até ao final do ano 2000.

No âmbito do Projecto Estatísticas do Turismo, foi realizado um estágio, para a Dra. Odete Lima, técnica da Direcção de Turismo de Cabo Verde, no sentido de dar continuidade à preparação do projecto.

Guiné Bissau

Face às perspectivas de regularização política e social na Guiné-Bissau, foi realizada uma reunião, aproveitando a estada em Lisboa do Dr. Serifo Monteiro, para análise das perspectivas de relançamento da actividade estatística e dos programas de cooperação.

Moçambique

No domínio do projecto de Regionalização e Descentralização da Actividade Estatística foi realizado um estágio, pelos técnicos do INE de Moçambique Dr. Bete Cordeiro e Sr. Jorge Chemane, que decorreu na Direcção Regional Norte do INE.

No âmbito do projecto de Disseminação e Gestão da Documentação e da Informação Estatísticas foi realizado um estágio, pelos técnicos do INE de Moçambique Sr. António Guimarães e Sr. Luís Rachid, que decorreu na Direcção Regional Centro do INE.

No domínio das Relações Externas, foi realizado um estágio, pela técnica do INE de Moçambique, Dra. Alda Rocha, tendo em vista o contacto com a organização e funcionamento dos serviços do INE.

Teve lugar Lisboa, a 8ª reunião da Comissão Coordenadora da Gestão do Acordo de Cooperação Estatística Luso - Moçambicano, na qual se estabeleceu as prioridades do programa de cooperação até ao final do ano 2000.

São Tomé e Príncipe

No âmbito do projecto Tecnologias de Informação e Informática foi realizado um estágio, para dois técnicos, Eng. Mário Coelho e Sr. Octávio Costa, que incidiu em operações de manutenção de hardware, técnicas de integração de softwares do Microsoft Office e operações de manutenção de fotocopiadoras.

Teve lugar Lisboa, a 12ª reunião da Comissão Coordenadora da Gestão do Acordo de Cooperação Estatística Luso - Santomense, tendo sido actualizado o programa de cooperação até ao final do ano 2000.

No âmbito do projecto Recenseamento Geral da População e Habitação, foi realizado um estágio, pelo Dr. Helder Salvaterra, técnico do INE de São Tomé e Príncipe, tendo em vista a continuação do programa de trabalho de preparação desta grande operação.

No âmbito do projecto de Difusão Estatística foi preparada a edição da publicação "Estatísticas Demográficas 1994-1996", a pedido extraordinário do INE de São Tomé e Príncipe.

b) Cooperação desenvolvida com os PECO, no quadro do Programa PHARE:

No âmbito do Programa PHARE de Assistência Técnica aos Países da Europa Central e Oriental, realizaram-se, durante o período mencionado em epígrafe, onze acções de cooperação.

Cinco dessas acções realizaram-se no âmbito de Projectos Piloto destinados aos países PHARE.

No âmbito do projecto piloto Estatísticas das Finanças Públicas, a Drª Ana Leal (DCCN) participou no segundo Workshop, realizado nos dias 27 a 29 de Setembro, na Roménia. O objectivo deste workshop consistiu na apresentação dos trabalhos realizados por cada país beneficiário no âmbito do projecto. No seguimento deste Workshop, a Drª Ana Leal realizou duas missões de assistência técnica aos países de que é responsável, com vista à preparação conjunta do relatório intermédio do projecto. As acções realizaram-se nos períodos de 15 a 16 de Novembro (Polónia) e 18 e 19 de Novembro (Eslovénia).

No quadro do projecto piloto Input Output Tables, a Drª Ana Barreno (DCCN), realizou duas missões de assistência técnica aos países da sua responsabilidade. A primeira acção teve lugar em Lubljana (Eslováquia), nos dias 27 a 29 de Outubro, e a segunda acção decorreu em Budapeste (Hungria), nos dias 1 a 3 de Novembro. O objectivo destas acções foi a inventariação dos trabalhos efectuados por cada um dos

organismos estatísticos com vista à preparação, por parte da técnica, de relatórios intermédios do projecto.

No âmbito da **Cooperação Bilateral**, realizaram-se seis acções de cooperação, das quais cinco com o *National Commission for Statistics* da Roménia, e uma com o *Central Statistical Office* da Hungria.

Roménia

A primeira acção, que teve lugar entre 8 a 10 de Setembro, consistiu num estágio ao INE, no âmbito do Projecto Apoio Institucional, financiado pelo PHARE Nacional. Esta acção teve como objectivo apresentar o Sistema Estatístico Português e analisar o actual sistema estatístico romeno com vista a apresentar sugestões de alterações para uma harmonização com os sistemas estatísticos da União Europeia. Este estágio contou com a participação do Dr. Ferreira da Cunha (consultor) e com o Dr. Pedro Dias (Director do DCCN).

A segunda acção, que revistiu também a forma de estágio, teve lugar na Direcção Regional Norte, no período de 15 a 26 de Novembro, e inseriu-se no quadro do Programa PHARE Nacional - Contas Nacionais/Regionais. Esta acção esteve a cargo da Dr^a Emília Saleiro e teve como objectivos a apresentação das opções metodológicas adoptadas pelo INE para a aplicação do Sistema Europeu de Contas 1995 e a realização de um estudo comparativo sobre as fontes de informação disponíveis.

A Dr^a Teresa Saraiva e a Dr^a Mariana Pereira (ambas do DEE), realizaram, na semana de 22 a 26 de Novembro, uma missão à Roménia, nas áreas das Estatísticas do Turismo e Transportes respectivamente. Estas acções tiveram como objectivo a análise do trabalho realizado pelo NCS e a definição dos trabalhos futuros, no âmbito das duas áreas estatísticas.

Por último, e no âmbito do Projecto Difusão do Programa PHARE Nacional, realizou-se um estágio no INE, no período de 17 a 21 de Dezembro, de técnicos da Roménia com o objectivo de finalizar a proposta de concurso internacional de aquisição de equipamento tipográfico. Esta acção serviu também para analisar o documento estratégico sobre a Difusão a apresentar pelo NCS, assim como para apresentar os produtos de difusão *on line* que estão a ser desenvolvidos pelo INE (esta componente esteve a cargo do Dr. Pedro Campos da DRN e as restantes componentes a cargo do Eng^o Pinto Martins do DDP).

Hungria

Com o organismo de estatística deste país apenas se realizou uma acção de cooperação, que ocorreu no INE no período de 20 a 24 de Setembro, e teve como objectivo a análise das redes informáticas existentes no Instituto, discussão das soluções informáticas desenvolvidas pelo INE para o problema do ano 2000 e a apresentação do Infoline, CyberINE e do Data Shop. Esta acção teve lugar no DSII (Eng^o Luís Silva), DRLVT (Eng^o Carlos Dias), DDP (Eng^o Pinto Martins) e DRC (Dr. Carlos Goulão).

c) Cooperação desenvolvida com os países do Mediterrâneo, no quadro do Programa MEDSTAT

Teve lugar em Atenas, entre 29 de Novembro a 1 de Dezembro, a 3ª reunião do Comité Director do Programa MEDSTAT, na qual este Instituto esteve representado pelo Dr. Ferreira da Cunha (Consultor). O objectivo desta reunião foi o de analisar a execução dos projectos no âmbito do referido Programa e a discussão de novos projectos de cooperação comuns e nacionais para o período pós-MEDSTAT.

d) Cooperação desenvolvida no âmbito do programa União Europeia – MERCOSUR

No âmbito deste programa realizaram-se, nos dias 22 e 23 de Setembro, em Montevidéu (Uruguais), as primeiras reuniões dos Grupos de Trabalho (GT) Nomenclaturas e Classificações e Estatísticas Sociais. A participação do INE foi assegurada pelo Dr. Saraiva Aguiar (DCCN) relativamente ao primeiro GT, e pelo Dr. Farrajota Leal para o GT Estatísticas Sociais. Estas reuniões contaram também com a participação de técnicos dos organismos de estatística dos países que constituem o MERCOSUR e tiveram como objectivo, por um lado, a apresentação do trabalho realizado em cada uma das áreas pelos vários países e, por outro lado, a apresentação, por parte dos peritos europeus, das metodologias e melhores práticas utilizadas no âmbito do sistema estatístico Europeu.

FUNDAMENTO, OBJECTO E ÂMBITO

O INE, consciente de como uma cultura estatística é essencial para a compreensão da maioria dos fenómenos do mundo actual, e da sua responsabilidade na divulgação do conhecimento estatístico, fazendo-o chegar ao maior número possível de leitores, tendo reconhecido a necessidade de dar um passo nesse sentido, passa a editar quadrimestralmente a presente Revista de Estatística destinada a divulgar:

- a) Numa perspectiva científica, artigos originais sobre temas especializados da estatística, tanto pura como aplicada, bem como sobre estudos e análises nos domínios económico, social e demográfico;
- b) Informações sobre actividades e projectos importantes no âmbito do Sistema Estatístico Nacional;
- c) Informações sobre congressos, seminários, colóquios e conferências de interesse estatístico ou afim;
- d) Informações sobre acções desenvolvidas pelo INE no âmbito da cooperação bilateral e multilateral.

Para tal, são adoptadas as seguintes formas de contribuição para publicação na Revista:

- Quanto aos artigos referidos em a), contribuições da iniciativa dos próprios autores e por convite do Conselho Editorial, pertencentes ou não ao INE;
- Quanto às informações referidas em b), c) e d), contribuições dos departamentos do INE.

As contribuições por iniciativa dos próprios autores serão objecto de avaliação de mérito científico pelo Conselho Editorial, que decidirá ou não pela respectiva publicação.

Para a elaboração e envio das contribuições para publicação na Revista são adoptadas as Normas de Apresentação de Manuscritos que figuram na última página.

Os autores dos artigos publicados, a que se refere a alínea a), receberão uma contribuição financeira paga pelo INE, de montante a fixar por despacho da Direcção mediante proposta do Director da Revista.

OS PONTOS DE VISTA EXPRESSOS PELOS AUTORES DOS ARTIGOS PUBLICADOS NA REVISTA

NÃO REFLECTEM NECESSARIAMENTE A POSIÇÃO OFICIAL DO INE.

FOUNDATION, SUBJECT MATTER AND SCOPE

INE is conscious of how statistical awareness is essential to the understanding of the majority of phenomena in the present world and is aware of its responsibility to disseminate statistical knowledge, making it available to the widest possible range of readers. INE has recognised the need to take a step in that direction and will begin publication of this *Statistical Review* three times yearly, designed to provide the following:

- a) Within a scientific perspective, original articles on specialised areas of statistics, both pure and applied, as well as studies and analyses within the sphere of economics, social issues and demographics;
- b) Information on activities and projects within the scope of the National Statistical System;
- c) Information on congresses, seminars and conferences of a statistical or related nature;
- d) Information on activities developed by INE within the scope of bilateral or multilateral co-operation;

The following approaches for contributing material for publication in the review have been adopted:

- In relation to the articles referred to in section a), contributions are made by the authors themselves and by invitation of the Editorial Committee, whether they are employees of INE or not;
- In relation to the information referred to in section b), c) and d); contributions are from departments of INE.

The Editorial Committee who has sole discretion in deciding whether or not the material will be published will assess the scientific merit of contributions made on the initiative of the authors themselves.

The preparation and delivery of material for publication in the Review are subject to the Rules for Submitting Manuscripts presented on the last page.

The authors of the published articles referred to in section a) will receive pecuniary compensation from INE in an amount to be determined by resolution of the Board on the recommendation of the Director of the Review.

THE VIEWPOINTS EXPRESSED BY THE AUTHORS OF THE ARTICLES PUBLISHED IN THE REVIEW

DO NOT NECESSARILY REFLECT THE OFFICIAL POSITION OF I.N.E.

NORMAS DE APRESENTAÇÃO DE MANUSCRITOS

Nos termos da alínea b) do nº. 3 do Artigo 5º do Regulamento da *Revista de Estatística* do Instituto Nacional de Estatística, o Conselho Editorial aprovou as seguintes **Normas de Apresentação de Manuscritos**:

1. Os originais dos artigos serão enviados ao Director da Revista pelos respectivos autores, devendo ser escritos em português e não terem sido ainda totalmente publicados, ou estar em processo de edição em qualquer outra publicação.
2. Poderão também ser apresentados artigos escritos em inglês, cabendo ao Director da Revista a decisão sobre a sua aceitação.
3. Quanto à *avaliação do mérito científico* dos artigos:
 - a) Os artigos apresentados por iniciativa dos respectivos autores serão submetidos à avaliação do mérito científico pelo Conselho Editorial, com garantia do anonimato tanto do autor como dos avaliadores;
 - b) Os autores receberão a informação sobre o resultado da avaliação num prazo máximo de trinta e cinco dias, com indicação, nos casos de avaliação positiva, do número da *Revista* em que serão publicados, e nos casos de avaliação negativa com a devolução do artigo apresentado e respectiva *disquette*, com indicação do(s) software(s) adicional(ais) eventualmente utilizado(s) na produção do documento original.
4. Os artigos aceites para publicação na *Revista de Estatística* serão igualmente divulgados no *site* do INE na *Internet*.
5. Os originais, com uma extensão não superior a trinta páginas, serão processados em *Word for Windows*, integralmente a preto e branco, e entregues em suporte papel acompanhado da respectiva *disquette*.
6. Na apresentação dos originais, os autores respeitarão ainda as seguintes normas:
 - 6.1. Quanto à *estrutura*:
 - a) O texto deve ser dactilografado em formato A₄, com utilização do tipo de letra *Times New Roman* - 11, e com as seguintes margens: *top*: 2,5 cm, *bottom*: 2 cm, *left*: 2,5 cm, *right*: 5 cm;
 - b) A primeira página conterà exclusivamente o título do artigo, bem como o nome, morada e telefone do autor, com indicação das funções exercidas e da instituição a que pertence, devendo, no caso de vários autores, ser indicado a quem deverá ser dirigida a correspondência da Revista;
 - c) A segunda página conterà, em português e inglês, unicamente o título e um resumo do artigo, com um máximo de cem palavras, seguido de um parágrafo com indicação de palavras-chave até ao limite de quinze;
 - d) Na terceira página começará o texto do artigo, sendo as suas eventuais secções ou capítulos numeradas sequencialmente;
 - 6.2. Quanto a *referências bibliográficas*:
 - a) Os autores eventualmente citados no texto do artigo serão indicados entre parênteses curvos pelo seu nome seguido da data da respectiva publicação e, se for caso disso, do número de página (p. ex.: Malinvaud, 1989, 23);
 - b) As referências bibliográficas serão listadas, por ordem alfabética dos apelidos dos respectivos autores, imediatamente a seguir ao final do texto, de acordo com a fórmula seguinte:

ANDERSON, C.W., and TURKMAN, K.F, (1995) "Sums and maxima of stationary sequences with heavy tailed distributions", *Sankhya*, Vol. 57, Series A, pp.1-10.

6.3. Quanto à *revisão de provas e publicação*:

- a) Uma vez aceite o artigo e antes da sua publicação, receberá o autor dois exemplares de provas para revisão, um dos quais será devolvido ao Director da Revista no prazo máximo de uma semana contado da data da sua recepção;
- b) Serão da responsabilidade dos respectivos autores as consequências de eventuais modificações da versão inicial aceite, bem como de atrasos na revisão das provas, que impossibilitem a publicação no número da Revista previsto, reservando-se o Conselho Editorial o direito de decidir a data da sua publicação futura;
- c) Uma vez publicado o artigo, o autor receberá vinte exemplares da sua versão impressa e um exemplar do respectivo número da *Revista*.

7 Para informações adicionais contactar o Secretariado de Redacção:

Eduarda Liliana Martins

Instituto Nacional de Estatística

Av^a. António José de Almeida, nº. 5 – 9^o.

1 000 Lisbon - Portugal

TEL.: +351 21 842 61 00 (3905)

FAX.: +351 21 842 63 66

E-MAIL: liliana.martins@ine.pt

RULES FOR SUBMITTING MANUSCRIPTS

Within the terms of sub-section a of no. 3 of Article 5 of the regulations of the *Statistical Review* of the National Statistical Institute (INE), the Editorial Committee has approved the following **Rules for Submitting Manuscripts**:

1. The original articles will be sent to the Review Director by the respective authors. They should be written in Portuguese, they should not have already been published in their entirety nor should they be in the process of being published in any other publication.
2. Articles may also be submitted in English to the Review Director who will decide whether to accept them.
3. In relation to the *evaluation of the scientific merit* of the articles:
 - a) The Editorial Committee will assess all articles submitted on the initiative of the respective authors on the basis of their scientific merit. The identity of both the author and the Committee members will be strictly confidential;
 - b) The authors will receive information regarding the results of the evaluation within a maximum period of thirty-five days. If the article is accepted, the Committee will indicate the issue number of the *Review* in which the article will be published. If the article is not accepted, it will be returned along with the respective diskette, with the information on the additional(s) software(s) eventually used in the production of the original document.
4. The articles accepted for publication in the *Statistical Review* will also be made public on the INE Internet site.
5. The original articles having no more than thirty pages must be processed in *Word for Windows*, completely at black and white, and they will be delivered in hard copy as well as on diskette.
6. With the presentation of the original articles, the authors must also respect the following rules:
 - 6.1 In relation to the *structure*:
 - a) The text shall be printed on A4 format paper utilising the font *Times New Roman* size 11 and with the following margins: top: 2.5 cm, bottom: 2 cm, left: 2,5 cm, right: 5 cm;
 - b) The first page shall contain only the title of the article as well as the name, address and telephone number of the author, indicating the position held and the institution that he/she belongs to. In the case of various authors, it is necessary to indicate the person to whom all correspondence received by the *Review* should be forwarded;
 - c) The second page shall contain only the title and a abstract of the article in Portuguese and English with the maximum of one hundred words followed by a paragraph indicating key words up to the limit of fifteen;
 - d) The third page will begin the text of the article with its respective sections or chapters sequentially numbered;
 - 6.2 Regarding *bibliographical references*:

- a) Authors who are cited in the text of the article shall be indicated in parentheses with their name followed by the date of the respective publication and, if necessary, the page number (ex.: Malinvaud, 1989, 23);
- b) All bibliographical references will be listed in alphabetical order by the surnames of the respective authors, immediately following the end of the text, as in the following example:

ANDERSON, C.W., and TURKMAN, K.F., (1995) "Sums and maxim of stationary sequences with heavy tailed distributions", *Sankhya*, Vol. 57, Series A, pp. 1-10.

6.3 Regarding *proof-reading and publication*:

- a) Once the article is accepted and prior to its publication, the author will receive two copies for review. One of these copies will be returned to the Director of the Review within a maximum period of one week from the date of its reception;
- b) The consequences of subsequent changes to the accepted first version are the responsibility of the respective authors as well as any delays in proof-reading that make its publication in the planned issue of the Review impossible. The Editorial Committee reserves the right to decide upon the date for future publication;
- c) Once the article is published, the author will receive twenty copies of his/her printed version and a copy of the respective issue of the *Review*.

7. For further information kindly contact the Editorial Secretary:

Euarda Liliana Martins
Instituto Nacional de Estatística
Av^a. António José de Almeida, n^o. 5 – 9^o.
1 000 Lisbon - Portugal
TEL.: +351 21 842 61 00 (3905)
FAX.: +351 21 842 63 66
E-MAIL: liliana.martins@ine.pt

LISTA DE PUBLICAÇÕES

Algumas Publicações Editadas pelo INE

Índice de Custo do Trabalho - Metodologia e 1º Resultados (1995 a 1º Trim. 1999)
 Nomenclaturas, Territoriais Designações e Códigos 1998
 Classificação Nacional de Bens e Serviços 1998

600\$00
 3.600\$00
 12.000\$00

ESTATÍSTICAS GERAIS

Anuário Estatístico de Portugal 1998	10.700\$00	8.600\$00	6
Boletim Mensal de Estatística 1999 (x 12)	2.400\$00	23.000\$00	1
Indicadores Urbanos do Continente 1999	5.100\$00		

POPULAÇÃO, AMBIENTE CONDIÇÕES SOCIAIS

Estatísticas da Cultura, Desporto e Recreio 1997	3.800\$00	3.000\$00	5
Série Estimativas Provisórias N° 28	3.900\$00		
Portugal Social 1991/1995	6.000\$00		
Estatísticas da Protecção Social 1997	2.160\$00	1.730\$00	5
Estatísticas da Saúde 1998	9.000\$00	7.200\$00	6
Estatísticas Demográficas 1998	6.600\$00	5.300\$00	6
Estatísticas do Ambiente 1997	3.000\$00	2.400\$00	5
Estatísticas do Emprego 1999 (Trimestral)	1.300\$00	4.200\$00	3

AGRICULTURA, SILVICULTURA E PESCA

Estatísticas da Pesca 1998	3.000\$00	2.400\$00	5
Inquérito às Plantações de Árvores de Fruto 1998	1.500\$00		
Estatísticas Agrícolas 1998	4.200\$00	3.400\$00	5
Pescas em Portugal 1996 - 1998	6.300\$00		
Contas Económicas da Agricultura 1998	1.500\$00		
Estado das Culturas e Previsão das Colheitas 1999	240\$00	2.300\$00	2

INDÚSTRIA, CONSTRUÇÃO E ENERGIA

Estatísticas da Construção de Edifícios 1997	2.120\$00	1.700\$00	5
Estatísticas da Produção Industrial 1997	4.300\$00	3.400\$00	5
Estatísticas das Empresas - Agricultura e Indústria 1997	2.700\$00	2.160\$00	5
Índices de Produção Industrial 1999	200\$00	2.200\$00	2
Estatísticas das Empresas - Construção 1997	1.500\$00	1.200\$00	5
Inquérito Mensal à Construção e Obras Públicas 1999	650\$00	6.200\$00	2
Índices de Preços na Produção Industrial 1999	430\$00	4.100\$00	2
Índices de Volume de Negócios, Emprego, Remunerações e Horas Trabalhadas na Indústria 1999	300\$00	3.600\$00	2
Inquérito Mensal à Indústria Transformadora 1999	720\$00	6.900\$00	2
Inquérito Mensal de Conjuntura Serviços Prestados às Empresas 1999	300\$00	2.900\$00	2

COMÉRCIO INTERNACIONAL

Comércio Internacional 1999	800\$00	8.500\$00	2
Estatísticas do Comércio Internacional 1998	8.100\$00	6.500\$00	6
Comércio Extra-Comunitário 1999	700\$00	6.700\$00	2

COMÉRCIO INTERNO, TURISMO E OUTROS SERVIÇOS

Estatísticas do Turismo 1998	4.700\$00	3.800\$00	6
Estatísticas dos Transportes e Comunicações 1998	6.300\$00	5.000\$00	6
Estatísticas das Empresas - Comércio e Outros Serviços 1997	9.000\$00	7.200\$00	6
Estatísticas do Transportes Rodoviários de Passageiros e de Mercadorias 1998	3.300\$00		
Gastos dos Estrangeiros não Residentes Residentes em Portugal 1997	1.220\$00		
Estabelecimentos Comerciais 1998	600\$00	720\$00	5
Índice do Volume de Negócios no Comércio a Retalho 1999	190\$00	1.800\$00	2
Inquérito Mensal de Conjuntura ao Comércio 1999	1.300\$00	12.500\$00	2

ECONOMIA E FINANÇAS

Estatísticas das Receitas Fiscais 1998	3.070\$00	2.460\$00	6
Empresas em Portugal 1990 - 1995	2.190\$00		
Estatísticas das Administrações Públicas 1997	2.300\$00	1.800\$00	5
Estatísticas Monetárias e Financeiras 1997	5.500\$00		
Sistema de Contas Integradas das Empresas 1995 - 1996	3.800\$00		
Índice de Preços no Consumidor 1999	1.400\$00	13.400\$00	2
Contas Nacionais 1995	2.070\$00		
Síntese Económica Mensal 1999	400\$00	4.600\$00	2

ESTATÍSTICAS REGIONAIS

Contas Regionais 1995	2.900\$00		
Retalho das Regiões 1998	5.000\$00		
Anuário Estatístico da Região Lisboa e Vale do Tejo 1998	6.000\$00		
Inventário Municipal da Região Lisboa e Vale do Tejo 1998	5.970\$00		
Inquérito ao Emprego Região de Lisboa e Vale do Tejo (NUTS III) 1999 (Semestral)	600\$00		
Índice de Preços no Consumidor - Região de Lisboa e Vale do Tejo 1999 (Mensal)	600\$00	5.800\$00	2
Anuário Estatístico da Região Algarve 1998	4.000\$00		
Inventário Municipal da Região Algarve 1998	4.600\$00		
Anuário Estatístico da Região ALENTEJO 1998	4.500\$00		
Inventário Municipal da Região ALENTEJO 1998	5.000\$00		
Anuário Estatístico da Região Centro 1998	6.000\$00		
Inventário Municipal da Região Centro 1998	6.000\$00		
Anuário Estatístico da Região Norte 1998	5.600\$00		

ESTUDOS

Revista de Estatística 1999 (quadrimestral)	2.500\$00	6.000\$00	7
---	-----------	-----------	---

* PORTES DE CORREIO

	PORTUGAL		EUROPA		RESTO DO MUNDO	
	Assin.	Anúlo	Assin.	Anúlo	Assin.	Anúlo
1	1.920\$00	160\$00	5.040\$00	420\$00	9.300\$00	775\$00
2	1.020\$00	85\$00	2.520\$00	210\$00	4.080\$00	340\$00
3	340\$00	85\$00	840\$00	210\$00	1.360\$00	340\$00
4	170\$00	85\$00	420\$00	210\$00	680\$00	340\$00
5	285\$00	285\$00	765\$00	765\$00	1.480\$00	1.480\$00
6	560\$00	560\$00	1.325\$00	1.325\$00	2.600\$00	2.600\$00
7	900\$00	300\$00	2.295\$00	765\$00	4.440\$00	1.480\$00



Nome _____ Data de nascimento: ____/____/____

Profissão/Função _____ Instituição/Empresa _____

Telef.: _____ Fax: _____

DESEJO RECEBER OS EXEMPLARES DA REVISTA DE ESTATÍSTICA:

Em casa Na Instituição/empresa

Morada para envio: _____

Localidade: _____ Código Postal: _____

Autorizo débito no cartão Visa ou Mastercard n.º:

Valor da transacção: 6. 900S00 Validade do cartão ____/____/____

 Junto cheque n.º _____ à ordem do INSTITUTO NACIONAL DE ESTATÍSTICA sobre o Banco _____

Data: ____/____/____ Assinatura: _____

OS DADOS RECEBIDOS SERÃO PROCESSADOS AUTOMATICAMENTE E DESTINAM-SE AOS ENVIOS RELACIONADOS COM A SUA ASSINATURA, RESPECTIVAS OPERAÇÕES ADMINISTRATIVAS E ESTATÍSTICAS, E À EVENTUAL APRESENTAÇÃO DE OUTROS PRODUTOS E SERVIÇOS DO INSTITUTO NACIONAL DE ESTATÍSTICA.

Nome _____ Data de nascimento: ____/____/____

Profissão/Função _____ Instituição/Empresa _____

Telef.: _____ Fax: _____

DESEJO RECEBER OS EXEMPLARES DA REVISTA DE ESTATÍSTICA:

Em casa Na Instituição/empresa

Morada para envio: _____

Localidade: _____ Código Postal: _____

Autorizo débito no cartão Visa ou Mastercard n.º:

Valor da transacção: 6. 300S00 Validade do cartão ____/____/____

 Junto cheque n.º _____ à ordem do INSTITUTO NACIONAL DE ESTATÍSTICA sobre o Banco _____

Data: ____/____/____ Assinatura: _____

OS DADOS RECEBIDOS SERÃO PROCESSADOS AUTOMATICAMENTE E DESTINAM-SE AOS ENVIOS RELACIONADOS COM A SUA ASSINATURA, RESPECTIVAS OPERAÇÕES ADMINISTRATIVAS E ESTATÍSTICAS, E À EVENTUAL APRESENTAÇÃO DE OUTROS PRODUTOS E SERVIÇOS DO INSTITUTO NACIONAL DE ESTATÍSTICA.



AUTORIZADO PELOS CTT
NO SERVIÇO NACIONAL

RSF
NÃO PRECISA DE SELO

INSTITUTO NACIONAL DE ESTATÍSTICA
SECÇÃO VENDA DE INFORMAÇÃO

Av. Antonio José de Almeida
1000-043 LISBOA

AUTORIZADO PELOS CTT
NO SERVIÇO NACIONAL

RSF
NÃO PRECISA DE SELO

INSTITUTO NACIONAL DE ESTATÍSTICA
SECÇÃO VENDA DE INFORMAÇÃO

Av. Antonio Jose de Almeida
1000-043 LISBOA





VA049903

