



INSTITUTO NACIONAL DE ESTATÍSTICA

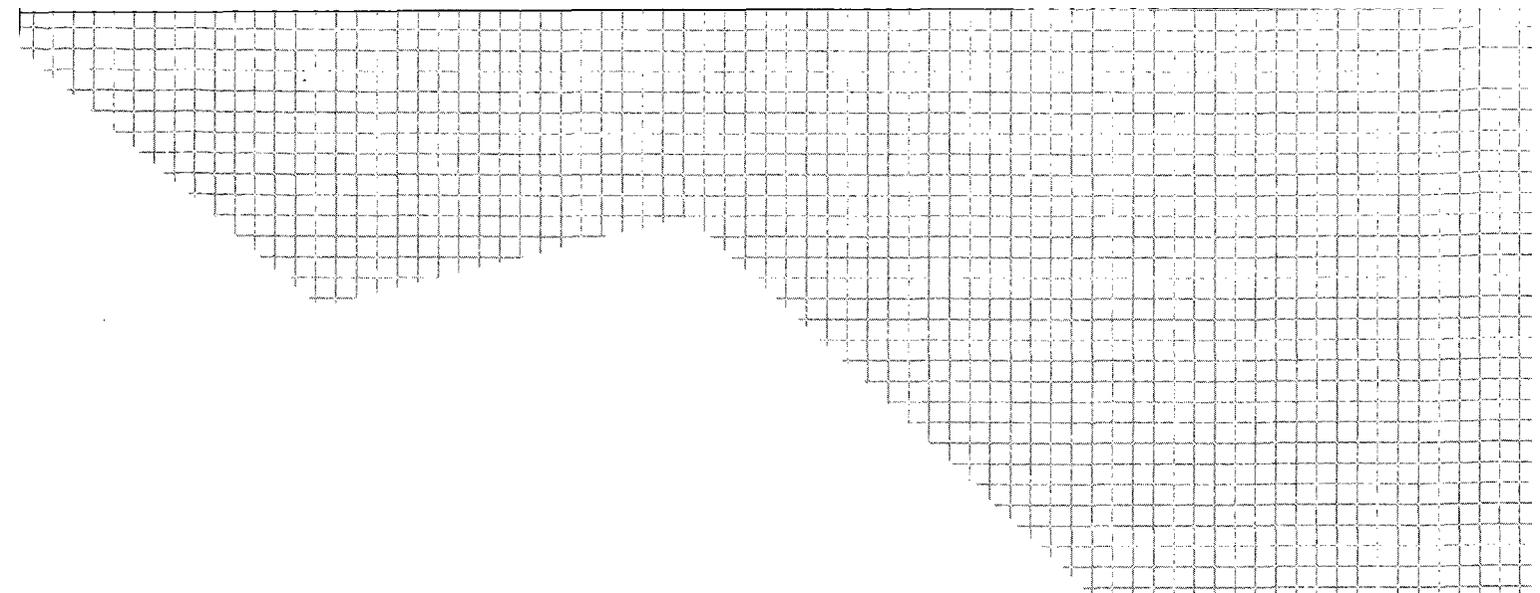
PORTUGAL

# REVISTA DE ESTATÍSTICA

## STATISTICAL REVIEW







# REVISTA DE ESTADÍSTICA

**STATISTICAL REVIEW**

**VOLUME II**

**2° QUADRIMESTRE DE 2002**

**ÍNDICE**

**INDEX**

- MENSAGEM DO DIRECTOR SOBRE O NOVO FORMATO DA REVISTA <i>DIRECTOR'S STATEMENT ON THE NEW JOURNAL FORMAT</i> .....	5
- ARTIGOS <i>ARTICLES</i> .....	11
ÍNDICES DE DESENVOLVIMENTO CONCELHIO <i>MUNICIPAL DEVELOPMENT INDEXES</i> Por/By: <i>Paulo A. Lucas da Fonseca</i> .....	13
UTILIZAÇÃO DE BASES DE MICRODADOS NA INVESTIGAÇÃO EM CIÊNCIAS SOCIAIS <i>THE UTILIZATION OF MICRO-DATABASES IN SOCIAL SCIENCES RESEARCH</i> Por/By: <i>Rosário Mauritti, Susana da Cruz Martins, Ana Simões Antunes, António Firmino da Costa</i> .....	47
DUALIDADES ENTRE ANÁLISE DE COVARIÂNCIA E ANÁLISE DE REGRESSÃO COM VARIÁVEIS DUMMY <i>DUALITIES BETWEEN ANCOVA AND REGRESSION WITH DUMMY VARIABLES</i> Por/By: <i>Patrícia Oom do Valle, Efigénio Rebelo</i> .....	65
AVALIAÇÃO DE UM MÓDULO DE FORMAÇÃO: ANÁLISE DAS EXPECTATIVAS E PERCEPÇÕES DOS FORMANDOS <i>EVALUATION OF A TRAINING MODULE: ANALYSES OF TRAINEES' EXPECTATIONS AND PERCEPTIONS</i> Por/By: <i>Maria José Conceição, Andrew B. Hill, Manuela Magalhães Hill</i> .....	87
$AM G \infty$ SYSTEMS COLLECTION WITH EXPONENTIAL BUSY PERIOD <i>UMA COLECÇÃO DE SISTEMAS <math>M/G/\infty</math> COM PERÍODO DE OCUPAÇÃO EXPONENCIAL</i> Por/By: <i>Manuel Alberto M. Ferreira</i> .....	111
- INFORMAÇÕES <i>INFORMATIONS</i> .....	123
CONGRESSOS, SEMINÁRIOS, COLÓQUIOS E CONFERÊNCIAS <i>CONGRESS, SEMINARS AND CONFERENCES</i> .....	125

FUNDAMENTO, OBJECTO E ÂMBITO DA REVISTA	
<i>FOUNDATION, SUBJECT MATTER AND SCOPE OF THE REVIEW</i> .....	133
NORMAS DE APRESENTAÇÃO DE ORIGINALS PARA A REVISTA	
<i>RULES FOR SUBMITTING ORIGINALS TO THE REVIEW</i> .....	135

**MENSAGEM DO DIRECTOR  
SOBRE O NOVO FORMATO DA REVISTA**

**DIRECTOR'S STATEMENT  
ON THE NEW JOURNAL FORMAT**

**VOLUME II**

**2° QUADRIMESTRE DE 2002**

---

## MENSAGEM DO DIRECTOR SOBRE O NOVO FORMATO DA REVISTA

---

Como os leitores terão certamente presente, no anterior Volume apresentei uma primeira informação sobre o projecto do Instituto Nacional de Estatística proceder à internacionalização da Revista de Estatística no próximo ano (2003), sendo oportuno relembrar os seguintes elementos caracterizadores do seu novo formato:

- O título será **REVSTAT - STATISTICAL JOURNAL**.
- A política editorial será centrada na publicação de artigos de investigação de alto nível nos domínios das Probabilidades e Estatística, com ênfase na originalidade e importância da investigação, sendo os artigos objecto de avaliação científica por pelo menos duas pessoas, uma do Corpo de Editores e outra externa.
- A língua de trabalho será exclusivamente o Inglês.
- A periodicidade da edição será quadrimestral, Março, Julho e Novembro, sendo o primeiro número publicado no final de Março de 2003.
- Serão publicados em média 6 artigos por cada número.
- **Corpo de Editores**, será constituído por:
  - **Editor-Chefe**, Prof.<sup>a</sup> *M. Ivette Gomes*, Faculdade de Ciências da Universidade de Lisboa, Portugal
  - **Co-Editor**, Prof.<sup>a</sup> *M. Antónia Amaral Turkman*, Faculdade de Ciências da Universidade de Lisboa, Portugal
  - **Editor Executivo**, Dr. *Ferreira da Cunha*, Consultor da Direcção do Instituto Nacional de Estatística, Portugal, e Director da actual Revista.
  - **Editores Associados:**
    - Prof. *Barry Arnold*, Universidade da Califórnia, Riverside, EUA.
    - Prof.<sup>a</sup> *H. Bacelar Nicolau*, Faculdade de Psicologia da Universidade de Lisboa, Portugal
    - Prof.<sup>a</sup> *Susie Bayarri*, Universidade de Valência, Espanha
    - Prof. *João Branco*, Instituto Superior Técnico de Lisboa, Portugal
    - Prof.<sup>a</sup> *M. Lucília Carvalho*, Faculdade de Ciências da Universidade de Lisboa, Portugal
    - Sir *David Cox*, Universidade de Oxford, Reino Unido
    - Prof. *Edwin Diday*, Universidade de Paris IX Dauphine, França
    - Prof. *Dani Gamerman*, Universidade Federal do Rio de Janeiro, Brasil
    - Prof.<sup>a</sup> *Marie Husková*, Universidade Carlos de Praga, República Checa
    - Prof. *Isaac Meilijson*, Universidade de Tel-Aviv, Israel
    - Prof.<sup>a</sup> *Nazaré Mendes-Lopes*, Faculdade de Ciências e Tecnologia da Universidade de Coimbra, Portugal
    - Prof. *António Pacheco*, Instituto Superior Técnico de Lisboa, Portugal
    - Prof. *Dinis Pestana*, Faculdade de Ciências da Universidade de Lisboa, Portugal

- Prof. *Ludger Rüschendorf*, Instituto de Matemática Estocástica de Freiburg, Alemanha
- Prof. *Gilbert Saporta*, Conservatório Nacional das Artes e Ofícios, Paris, França
- Prof. *Jef Teugels*, Universidade Católica de Lovaina, Bélgica
- Prof. *Radu Theodorescu*, Universidade de Laval, Sainte-Foy, QC, Canadá

- **Secretária:**

- *Liliana Martins*, Instituto Nacional de Estatística, Lisboa, Portugal ([liliana.martins@ine.pt](mailto:liliana.martins@ine.pt))

- Os artigos deverão ser escritos em Inglês e poderão ser submetidos das seguintes três maneiras:
  - Enviando 3 cópias em papel ao Editor Executivo, que estará em contacto permanente com o Editor-Chefe e o Co-Editor,
  - Enviando 2 cópias em papel ao Editor Executivo e uma cópia a um dos dois Editores ou dos Editores Associados, cuja opinião o autor queira que seja tomada em linha de conta,
  - Enviando 1 cópia em papel ao Editor Executivo, conjuntamente com o correspondente ficheiro PDF ou Postscript para o E-mail: [revstat@fc.ul.pt](mailto:revstat@fc.ul.pt).
- Os manuscritos devem ser escritos numa só página de folha A4 em duplo espaço, com uma margem do lado esquerdo de pelo menos 3cm, e não podem ter mais de 30 páginas. A primeira página deverá incluir o nome, instituição e endereço do(s) autor(es), e um resumo com o máximo de 100 palavras, seguido das palavras-chave num máximo de 6, e a classificação do assunto AMS 2000. Os autores deverão escrever a versão final dos artigos utilizando o LaTeX, no estilo da REVSTAT. Este estilo (REVSTAT.sty), exemplos de ficheiros (REVSTAT.tex), e instruções adicionais aos autores poderão ser obtidas **a partir do dia 9 de Setembro deste ano através do endereço web do INE**: <http://www.ine.pt/revstat>.
- A submissão de artigos significa que contêm um trabalho original ainda não publicado nem em vias de ser publicado sob qualquer forma algures.

**Os autores que queiram submeter artigos para publicação na REVSTAT poderão começar a fazê-lo a partir do dia 1 de Outubro deste ano.**

Quanto ao preço de aquisição da REVSTAT, a assinatura anual será de 24€, e o Volume avulso será 10€.

Informações complementares sobre esta iniciativa, direccionadas aos potenciais autores de artigos, serão dadas no próximo número da *Revista de Estatística* a ser editado na 1ª quinzena de Janeiro do próximo ano.

---

## DIRECTOR'S STATEMENT ON THE NEW JOURNAL FORMAT

---

As readers will certainly recall, in the last Volume I have presented a first information on the National Statistical Institute's project to internationalise the *Revista de Estatística* next year (2003), being opportune to recall the main features characterising its format:

- The title will be **REVSTAT – STATISTICAL JOURNAL**.
- The editorial policy will focus on publishing research articles at the highest level in the domains of Probability and Statistics, with emphasis on the originality and importance of the research. All articles will undergo scientific evaluation by at least two persons, one from the Editorial Board and another, external.
- The only working language allowed will be English.
- The edition will be published every four months, in March, July and November, with the first issue coming out at the end of March 2003.
- On average, 6 articles will be published per issue.
- The **Editorial Board** will comprise:
  - **Editor-in-Chief**, Prof. *M. Ivette Gomes*, Faculty of Science, University of Lisbon, Portugal
  - **Co-Editor**, Prof. *M. Antónia Amaral Turkman*, Faculty of Science, University of Lisbon, Portugal
  - **Executive Editor**, Dr. *Ferreira da Cunha*, Consultant to the Board of Directors of the National Statistical Institute, Portugal, and Director of the *Revista de Estatística*.
  - **Associate Editors**:
    - Prof. Barry *Arnold*, University of California, Riverside, USA.
    - Prof. H. Bacelar Nicolau, Faculty of Psychology, University of Lisbon, Portugal
    - Prof. Susie Bayarri, University of Valencia, Spain
    - Prof. João Branco, Higher Technical Institute of Lisbon, Portugal
    - Prof. M. Lucília Carvalho, Faculty of Science, University of Lisbon, Portugal
    - Sir David Cox, University of Oxford, United Kingdom
    - Prof. Edwin Diday, University of Paris IX Dauphine, France
    - Prof. Dani Gamerman, Federal University of Rio de Janeiro, Brazil
    - Prof. Marie Husková, Charles University of Prague, Czech Republic
    - Prof. Isaac Meilijson, University of Tel Aviv, Israel
    - Prof. Nazaré Mendes-Lopes, Faculty of Science and Technology, University of Coimbra, Portugal
    - Prof. António Pacheco, Higher Technical Institute of Lisbon, Portugal
    - Prof. Dinis Pestana, Faculty of Science, University of Lisbon, Portugal

- Prof. Ludger Rüschendorf, Institute of Stochastic Mathematics, Freiburg, Germany
  - Prof. Gilbert Saporta, National Conservatory of Arts and Crafts, Paris, France
  - Prof. Jef Teugels, Catholic University of Leuven, Belgium
  - Prof. Radu Theodorescu, University of Laval, Sainte-Foy, QC, Canada
- **Secretary:**
- *Liliana Martins*, National Statistical Institute, Lisbon, Portugal ([liliana.martins@ine.pt](mailto:liliana.martins@ine.pt))
- The articles must be written in English and may be submitted in the following three different ways:
- By sending 3 paper copies to the Executive Editor who will be in permanent contact with the Editor-in-Chief and the Co-Editor,
  - By sending 2 paper copies to the Executive Editor and one copy to one of the two Editors or Associate Editors, whose opinion the author wants to be taken into account.
  - By sending 1 paper copy to the Executive Editor, together with the corresponding PDF or PostScript file to the following e-mail address: [revstat@fc.ul.pt](mailto:revstat@fc.ul.pt).
- Manuscripts submitted must be typed on one side using A4-size paper, double-spacing, and a left margin of at least 3cm, and must not contain more than 30 pages. The first page must include the name, institution and address of the author(s) and a short abstract of not more than 100 words, followed by a maximum of 6 key words and the AMS 2000 subject classification. The authors must write the final version of the articles using LaTeX, in REVSTAT style. This style (REVSTAT.sty), examples file (REVSTAT.tex) and additional instructions for the authors may be obtained **from 9<sup>th</sup> September 2002 through the following INE Web address:** <http://www.ine.pt/revstat>.
- The submission of an article means that it contains original work that has not been nor is about to be published elsewhere in any form.

Within this new context, **authors wishing to submit articles for publication in REVSTAT may begin to do so from 1<sup>st</sup> October 2002.**

Concerning the price of **REVSTAT**, the annual subscription will be 24€, and the single issue will be 10€.

If necessary, further information on this initiative will be provided in the next issue of the *Revista de Estatística*, to be published in January 2003.

# ARTIGOS

**VOLUME II**

**2° QUADRIMESTRE DE 2002**

# Índices de Desenvolvimento Concelhio

Autor:  
Paulo A. Lucas da Fonseca

**VOLUME II**

**2º QUADRIMESTRE DE 2002**

---

---

## ÍNDICES DE DESENVOLVIMENTO CONCELHIO<sup>1</sup>

---

---

---

---

## MUNICIPAL DEVELOPMENT INDEXES

---

---

Autor: Paulo A. Lucas da Fonseca  
- Técnico Superior de Estatística – Instituição: Instituto Nacional de  
Estatística – Direcção Regional do Alentejo

### RESUMO:

- Este artigo tem como objectivo quantificar o actual nível de desenvolvimento económico e social das regiões e concelhos de Portugal Continental. Apresentam-se cinco índices parciais de desenvolvimento (demografia, saúde e assistência social, educação e cultura, rendimento, emprego e actividade económica), apoiados em aplicações de análise factorial de componentes principais. Adicionalmente, calcula-se um índice global de desenvolvimento, que sintetiza a informação dos cinco primeiros. Explica-se a metodologia utilizada e analisam-se os resultados que se reproduzem com o apoio de cartogramas e quadros.

### PALAVRAS-CHAVE:

- *Desenvolvimento, demografia, saúde e assistência social, educação e cultura, rendimento, emprego e actividade económica, índice, análise factorial de componentes principais, NUTS, concelho.*

### ABSTRACT:

- This paper aims at quantifying the level of economic and social development of the Portuguese Mainland regions and municipalities. With this aim, five partial development indexes (demography, health and social security, education and culture, income, employment and economic activity), supported by factor analysis (principal components analysis), are presented. In addition, a global development index summarising the previous five indexes is also presented. The approach employed to develop the indexes is explained and results are analysed with the support of maps and tables.

### KEY-WORDS:

- *Development, demography, health and social security, education and culture, income, employment and economic activity, index, principal components factor analysis, NUTS, municipality.*

---

<sup>1</sup> A composição gráfica dos mapas e quadros contaram com a colaboração de Maria José Patrocínio Silva, da Direcção Regional do Alentejo do Instituto Nacional de Estatística.

**VOLUME II**

**2° QUADRIMESTRE DE 2002**

## 1. INTRODUÇÃO

Colocar a informação ao serviço do desenvolvimento para o medir e monitorizar é, porventura, o fim último da produção e difusão de informação estatística. Com base nos resultados assim alcançados é possível planear e concretizar políticas mais adequadas às diferentes realidades económicas e sociais.

Quantificar o desenvolvimento não é uma tarefa fácil. As primeiras dificuldades resultam da sua plasticidade conceptual. Numa acepção abrangente pode considerar-se o desenvolvimento como um processo multifacetado com propósitos associados à acessibilidade de oportunidades e à fruição de meios orientados para o progresso e para uma vivência saudável. Numa óptica sectorial atende às mais variadas dimensões da vida humana. Numa óptica regional pressupõe distribuição espacial. Em qualquer acepção o desenvolvimento é sempre algo inacabado e a sua medição relativa.

Para além do plano conceptual, outras dificuldades há – associadas à disponibilidade de informação estatística – que entram a tarefa de aferição do desenvolvimento. Por vezes, não é tanto a falta de fontes e informação que incomoda como a sua abundância pouco sistematizada, criando dificuldades de selecção, recolha e tratamento analítico.

Seja como for, e também por isso, a estatística deve servir aquele objectivo, ainda que só o possa fazer por aproximações da realidade. Apesar de todas as limitações inerentes ao processo de redução dos fenómenos humanos a números, o conhecimento sistematizado que resulta desse processo é a manifestação do labor estatístico e o seu contributo inestimável para o próprio desenvolvimento.

O trabalho que se apresenta neste artigo quantifica o actual nível de desenvolvimento económico e social das regiões e dos concelhos de Portugal Continental. Os resultados são apresentados na forma de índices parciais de desenvolvimento, tantos quantas as áreas consideradas na análise – demografia, saúde e assistência social, educação e cultura, rendimento, emprego e actividade económica – e num índice global de desenvolvimento, que sintetiza a informação dos cinco primeiros. Estes índices são medidas resumidas de múltiplos indicadores económicos e sociais, extraídos de fontes diversas. Sem perder de vista a riqueza de informação residente nos indicadores singulares, os valores dos índices calculados resultam e reflectem o que há de essencial na multiplicidade dos fenómenos regionais e concelhos analisados, facilitando a sua leitura. No plano político serve-se o delineamento de prioridades e a tomada de decisão.

Para a definição das áreas de análise e dos indicadores de base as opções sujeitaram-se à informação disponível ao nível concelhio. Atendidas as limitações existentes, o modelo global reflecte as vertentes económicas e sociais que se consideraram primordiais na análise do desenvolvimento. Na sua concepção procurou-se uma simbiose de componentes várias, sem se atribuir primazia a nenhuma delas. O modelo global assenta, portanto, em pressupostos teóricos, é condicionado pela informação disponível e precede a análise estatística. Não retrata as características de um qualquer modelo obtido a partir de *outputs* estatísticos, ao contrário, mede o

desempenho regional e concelhio em fileiras com significado económico e social, previamente segmentadas <sup>2</sup>.

A selecção dos indicadores foi auxiliada por testes estatísticos, visando, por ajustes sucessivos, a consistência interna dos modelos parciais e a adequação entre os modelos e a metodologia adoptada. Na fase inicial do trabalho projectou-se uma base com várias dezenas de indicadores, fixando-se posteriormente a análise num conjunto de 39 indicadores (ver quadro 1).

A exposição do trabalho desenvolvido prossegue com uma nota que explica, sucintamente, a metodologia utilizada e com uma breve análise dos principais resultados obtidos em cada índice e para cada região. Os valores dos índices são reproduzidos no final do artigo, com o apoio de cartogramas e quadros.

## 2. NOTA METODOLÓGICA

Os índices foram calculados com recurso à *análise factorial de componentes principais* (ACP) <sup>3</sup>, apoiada no programa informático SPSS (Statistical Package for the Social Sciences) <sup>4</sup>. Esta análise incidiu sobre os 275 concelhos existentes no território de Portugal Continental em 1998, antes da criação dos municípios de Odivelas, Trofa e Vizela, oficializada no último quadrimestre desse ano (ver INE, 1999). Estes três concelhos não foram incluídos na análise, uma vez que não há informação sobre eles, referenciada a 1998, para a maioria dos indicadores de base utilizados. Desta situação resulta que os valores de base respeitantes a estes municípios estejam reportados aos respectivos concelhos de origem: Loures no caso de Odivelas; Santo Tirso no caso de Trofa; Felgueiras, Guimarães e Lousada no caso de Vizela.

No plano regional os valores dos índices reportam-se ao agrupamento dos concelhos segundo a Nomenclatura de Unidades Territoriais para fins Estatísticos, que divide o território continental em 5 NUTS de nível II e em 28 NUTS de nível III (ver

<sup>2</sup> Neste particular a metodologia adoptada assemelha-se à de Brandão, Pires e Portugal (1998).

<sup>3</sup> O principal atributo da ACP é o de reduzir uma grande quantidade de informação inicial (os indicadores originais) a um pequeno conjunto de variáveis (as componentes principais), representativas e sem perdas significativas de informação dos modelos originais (Reis, 1997, 253). O principal objectivo deste trabalho é precisamente esse: resumir em poucas variáveis e tornar facilmente perceptível o que se expressa, normalmente, numa multiplicidade de factores e dificilmente se apreende na totalidade. Genericamente, o modelo das componentes principais (ou factores) descreve-se da seguinte maneira:

$$F_m = x_1 w_{m1} + x_2 w_{m2} + \dots + x_n w_{mn}, \text{ em que:}$$

$F_m$  = factor derivado;

$x_{1..n}$  = valores das  $n$  variáveis iniciais;

$w_{m,1..n}$  = contribuições das  $n$  variáveis iniciais para a formação do factor  $F_m$ .

<sup>4</sup> Para uma utilização documentada da ACP a partir do SPSS em ambiente Windows, consultar, por exemplo: Pestana e Gageiro (2000), Norusis e SPSS INC (1994), Kinnear e Gray (1994). Sobre a ACP ver também: Reis (1997), Hair *et al* (1992), Dillon e Goldstein (1984), Bouroche (1989).

INE, 1999). Relativamente ao concelho de Gavião, a sua passagem administrativa da NUTS II de Lisboa e Vale do Tejo para a NUTS II do Alentejo, datada de Agosto de 1999, é posterior à data de referência dos indicadores de base. Como tal, e por questões também associadas à comparabilidade com outros dados, optou-se por manter este concelho na primeira daquelas NUTS.

Quadro 1

Indicadores Originais	Ano	Fonte
<b>Demografia</b>		
População residente / 1000 habitantes no Continente	1998	INE
Crescimento demográfico natural / 1000 habitantes	1998	INE
Crescimento demográfico migratório / 1000 habitantes	1998	INE
População residente c/ 0-14 anos / 100 habitantes c/ 65+ anos	1998	INE
População residente c/ 15-64 anos / 100 habitantes c/ 0-14 e 65 + anos	1998	INE
Alojamentos clássicos / 1000 alojamentos clássicos do Continente	1998	INE
<b>Saúde e Assistência Social</b>		
Médicos em hospitais, centros de saúde e postos médicos / 10000 habitantes	1998	INE
Pessoal enfermagem em hospitais, c. saúde e postos médicos / 10000 habitantes	1998	INE
Camas de hospitais e centros de saúde / 10000 habitantes	1998	INE
Farmácias e postos de medicamentos / 10000 km <sup>2</sup>	1998	INE
Índice sintético saúde (hospitais, c. saúde, extensões, postos médicos) / 10000 km <sup>2</sup>	1998	INE
Doentes saídos de hospitais e c. saúde c/ 10-64 anos / 100 habitantes c/ 10-64 anos	1998	INE
Capacidade equipamentos da Seg. Social p/ infância e juventude / 1000 habitantes < 15 anos	1998	IGFSS
Capacidade equipamentos da Seg. Social p/ 3ª idade / 1000 habitantes c/ 65 + anos	1998	IGFSS
<b>Educação e Cultura</b>		
Índice sintético educação (escolas de ensino básico e secundário) / 10000 km <sup>2</sup>	1998/99	DAPP
Professores do ensino básico e secundário / 10000 habitantes	1998/99	DAPP
Alunos do ensino básico e secundário / 1000 habitantes	1998/99	DAPP
Alunos do ensino superior / 10000 habitantes	1998/99	DAPP
Livros existentes em bibliotecas / 100 habitantes	1998	DAPP
Sessões de espectáculos públicos / 10000 habitantes	1998	DAPP
Espectadores de espectáculos públicos / 1000 habitantes	1998	DAPP
Despesas das Câmaras Municipais em cultura / 100 habitantes	1997, 98	DAPP
<b>Rendimento</b>		
Rendimento bruto declarado em IRS / agregados passivos de IRS	1998	DGI
Remuneração média mensal dos trabalhadores por conta de outrem	1998	IGFSS
Pensões de invalidez, velhice e sobrevivência (activos) / pensionistas activos	1998	INE
Juros de depósitos bancários de particulares / habitantes	1998	INE
Crédito hipotecário concedido a particulares / habitantes	1998	INE
Parque automóvel de ligeiros / 1000 habitantes	1998	ISP
<b>Emprego e Actividade Económica</b>		
Pessoal ao serviço nos estabelecimentos / 1000 habitantes	1998	IGFSS
Agregados passivos de IRS / 1000 habitantes	1998	DGI
População residente c/ 15-64 anos / desempregados registados	1998	INE/IEFP
Quadros superiores e médios / 1000 pessoas ao serviço em estabelecimentos	1998	IGFSS
Trabalhadores c/ hab. académ. superior / 10000 trabalhadores dos estabelecimentos	1998	IGFSS
Empresas sediadas / 1000 habitantes	1998	INE
Sociedades sediadas constituídas durante o ano / 10000 empresas sediadas	1998	INE
Estabelecimentos / 10000 habitantes	1998	IGFSS
Volume de negócios declarado em IRC / empresas declarantes	1998	DGI
Bancos e Seguradoras / 100000 habitantes	1998	INE
Pessoal ao serviço nos Bancos / 10000 habitantes	1998	INE

Notas:

Os índices sintéticos de saúde e de educação ponderam os estabelecimentos pela respectiva dimensão média. Esta dimensão média foi calculada a partir do número médio de médicos e enfermeiros e do número médio de professores e alunos de cada estabelecimento do Continente.

No indicador das despesas das Câmaras Municipais em cultura utilizou-se a média aritmética dos valores de dois anos de referência para minimizar os efeitos das oscilações anuais.

A aplicação da ACP foi executada de uma forma parcelar, tantas vezes quantos os modelos em análise. Os valores de cada um dos cinco índices parciais de desenvolvimento resultam, portanto, de cada uma das aplicações factoriais. O índice global de desenvolvimento provém da média aritmética dos valores dos cinco índices parciais.

A aplicabilidade da análise factorial por componentes principais a um determinado conjunto de dados depende, em grande medida, do grau de coesão existente entre as variáveis originais. Quanto menor for a correlação entre as variáveis que formam o modelo original mais difícil se torna reduzir a informação original a um pequeno número de componentes principais.

No caso apresentado, a consistência interna dos cinco modelos parciais e a aplicabilidade metodológica da ACP a estes modelos foram previamente testadas pelos resultados obtidos em diferentes fases da análise estatística. A leitura da matriz de correlações parciais dos indicadores originais e da matriz de *comunalidades*<sup>5</sup> de cada modelo parcial permitiu que se chegasse, por ajustes sucessivos na selecção dos indicadores, a modelos coesos, com coeficientes de correlação suficientemente fortes e estatisticamente significativos. Além disso, os testes de *Kaiser-Meyer-Olkin* e de *Bartlett*<sup>6</sup> apresentam valores satisfatórios em todas as aplicações factoriais, comprovando a adequação entre os modelos criados e a metodologia adoptada.

Na aplicação da ACP foram retidas, em cada modelo parcial, as componentes principais mais importantes, considerando-se para o efeito as que apresentaram um valor-próprio superior à unidade (isto é, superior ao valor-próprio de qualquer indicador original)<sup>7</sup>. Posteriormente, as componentes retidas foram submetidas a uma rotação ortogonal, que permitiu maximizar a relação de cada indicador original numa só componente principal e, desta forma, uma melhor interpretação do respectivo conteúdo. Experimentadas várias fórmulas, optou-se pela rotação *quartimax*.

A interpretação das componentes sujeitas a rotação, por análise dos respectivos *loadings*<sup>8</sup>, permitiu concluir pela conveniência em aproveitar-se a informação de todos os factores retidos. Com efeito, nos modelos factoriais analisados todas as componentes retidas estão fortemente correlacionadas com pelo menos um indicador original e cuja influência não deve ser negligenciada. Os valores concelhios (*scores*) em cada índice parcial resultam, desta forma, da média dos *scores* obtidos em cada componente retida, ponderada pela contribuição relativa de cada componente na

---

<sup>5</sup> As comunalidades obtidas posteriormente à extracção dos factores representam a variância dos indicadores que é explicada pelo conjunto dos factores e variam entre 0 e 1. No caso apresentado a grande maioria das comunalidades atinge ou ultrapassa o valor 0,7, o que denota modelos consistentes.

<sup>6</sup> O valor do teste de KMO varia entre 0 e 1. Neste estudo todos os modelos apresentam valores próximos ou superiores a 0,7, viabilizando a utilização da ACP. O teste de esfericidade de Bartlett questiona a hipótese nula segundo a qual a matriz em análise equivale a uma matriz-identidade. Nos modelos apresentados os valores de significância estatística permitem rejeitar sempre a hipótese nula.

<sup>7</sup> Na ACP há tantas componentes quantos os indicadores originais e só se devem considerar as mais importantes, ou seja as que explicam a maior parte da variância global do modelo. Quando obtidas a partir de matrizes de correlações, como foi o caso, o valor-próprio médio das componentes é igual à unidade, pelo que normalmente se retêm as que superam esse valor.

<sup>8</sup> Os *loadings* medem o grau de associação entre os indicadores originais e as componentes principais e expressam-se em coeficientes de correlação.

explicação do modelo <sup>9</sup>. Neste cálculo, os *scores* concelhios de cada componente traduzem o somatório dos produtos obtidos entre os valores *standardizados* dos indicadores originais e os coeficientes demandados à ACP pelo método de regressão para cada relação indicador-factor.

Os valores originais dos *scores* concelhios produzidos pela ACP são valores *standardizados*, com média nula e desvio-padrão unitário. Não obstante, os valores dos índices aqui reproduzidos foram transformados, aritmeticamente, em escalas de medida que assumem uma média 100 e um coeficiente de variação igual ao de um dos indicadores originais. Para esse efeito, o indicador original foi seleccionado de entre os que, em cada modelo, apresentavam maior correlação com o índice concelho ponderado (ver quadro 2) <sup>10</sup>.

O índice global de desenvolvimento concelho resulta da média aritmética dos cinco índices parciais ponderados. Assegurada a consistência interna dos modelos parciais, esse procedimento atribui a cada um a mesma importância no cálculo do índice global. Neste caso o procedimento para a mudança de escala é semelhante ao dos índices parciais, seleccionando-se o coeficiente de variação de entre todos os indicadores que compõem a base original de indicadores.

Os valores dos índices para os agrupamentos concelhios considerados – NUTS de nível I, II e III – são calculados mediante a ponderação dos índices dos concelhos pela respectiva população residente. Este expediente reproduz a importância relativa das parcelas e deixa em aberto a possibilidade de se calcular os índices para qualquer agregado concelho, sem necessidade de se recorrer à análise factorial.

**Quadro 2**

Coeficientes de Variação Atribuídos aos Índices		
Índices	Indicadores Originais	C.V. (%)
Desenvolvimento	Rendimento bruto declarado em IRS / agregados passivos de IRS	19,79
Demografia	População residente c/ 15-64 anos / 100 habitantes c/ 0-14 e 65 + anos	17,07
Saúde e Assistência Social	Capacidade equipamentos Seg. Social p/ inf. e juventude / 1000 hab. < 15 anos	63,35
Educação e Cultura	Professores do ensino básico e secundário / 10000 habitantes	20,61
Rendimento	Rendimento bruto declarado em IRS / agregados passivos de IRS	19,79
Emprego e Activ. Económica	Agregados passivos de IRS / 1000 habitantes	25,07

<sup>9</sup> Processo de ponderação idêntico é realizado em INE (2000).

<sup>10</sup> Processo de transformação aritmética idêntico é realizado em INE (1993).

---

### 3. ANÁLISE DE DADOS

---

Os índices de desenvolvimento apresentam em todos os modelos analisados distribuições regionais e concelhias assimétricas. Em todos os índices a maior parte das NUTS de nível II e III e a grande maioria dos concelhos posicionam-se abaixo da média do Continente, revelando as diferenciações espaciais do território continental no que concerne às características económicas e sociais analisadas.

Sobre essas diferenciações, os traços gerais que ressaltam dos vários índices caracterizam-se por uma preponderância de algumas zonas litorais face a zonas mais interiores do território continental e por uma predominância generalizada dos concelhos sede de distrito em relação aos demais. Esta última situação tem a particularidade de relevar não só as assimetrias infra-regionais, mas também o benefício relativo daí decorrente para algumas zonas do interior no que se refere aos índices de desenvolvimento logrados. Em alguns casos, e em particular nalgumas zonas menos dotadas do interior, percebe-se a importância dos concelhos capitais de distrito para os resultados globais obtidos pelas suas áreas envolventes.

---

#### 3.1. DESENVOLVIMENTO(S): ÍNDICES SOCIAIS E ECONÓMICOS

---

O índice global de desenvolvimento concelhio agrega a informação de todos os índices parciais, atenua assimetrias, mas demonstra as diferenças existentes. Ao nível das NUTS II constata-se que Lisboa e Vale do Tejo apresenta o valor mais elevado, 116,0, ou seja 16% acima da média do Continente. Todas as outras NUTS II registam valores inferiores a esta média, com destaque negativo para o índice do Alentejo, cifrado em 82,8, isto é, 17,2% abaixo do valor continental.

A análise por NUTS III permite verificar que a posição lograda por Lisboa e Vale do Tejo se deve não só ao valor da Grande Lisboa (o maior de entre todas as NUTS III do Continente), mas também ao menor afastamento negativo médio das suas restantes NUTS III face à média 100. O Norte e o Centro têm, igualmente, uma NUTS III (Grande Porto e Baixo Mondego, respectivamente) acima desse limiar, mas o afastamento médio negativo das restantes NUTS III ao valor do Continente é substancial. No Alentejo a homogeneização processa-se por baixo, não havendo nenhuma NUTS III que logre aproximar-se significativamente daquela média.

A análise do território continental por concelhos no índice global de desenvolvimento denota uma diferenciação entre as zonas do litoral e do interior. De uma forma geral, destaca-se pela positiva uma estreita faixa infra-regional no litoral, delimitada aproximadamente pelos concelhos de Braga e Setúbal, com sequência sem contiguidade na costa algarvia. Os concelhos de Lisboa e do Porto posicionam-se no topo deste grupo, o que demonstra a sua importância nas respectivas áreas envolventes, mas também no país. A este litoral mais desenvolvido juntam-se, igualmente, alguns concelhos sede de distrito do interior, com particular destaque para os casos de Viseu, Guarda, Évora e Beja, com valores acima da média continental.

Situação diferente é a do restante território que, com maior incidência em algumas zonas interiores do Norte, Centro e Alentejo, denota valores inferiores e genericamente distantes dessa média. Entre os resultados destes dois grupos extremos, constata-se uma preponderância de alguns concelhos capitais de distrito: Bragança, Vila Real, Castelo Branco e Portalegre, no interior, e Viana do Castelo, no litoral, são paradigmas deste grupo concelhio intermédio.

O índice de demografia tem por base indicadores de volume populacional e habitacional, crescimento e vitalidade demográfica. Neste caso, Lisboa e Vale do Tejo e Norte são as únicas NUTS II com índices superiores à média do Continente, 110,6 e 101,9, respectivamente. Na primeira há duas NUTS III que superam a média continental – Grande Lisboa e Península de Setúbal –, na segunda há cinco NUTS III que o conseguem – Cávado, Ave, Grande Porto, Tâmega e Entre Douro e Vouga. O Alentejo, por seu lado, denota uma grande debilidade demográfica, com o mais baixo índice (75,7) registado entre as NUTS II. Nesta região, a NUTS III com o resultado mais favorável é o Alentejo Litoral, ainda assim 21,5% abaixo da média 100. No Centro, com um valor global de 85,8, não há nenhuma NUTS III que atinja aquela média, ao passo que no Algarve o resultado é ainda um pouco mais baixo (84,8).

Os valores desagregados por concelhos revelam uma preponderância de alguns municípios das duas áreas metropolitanas. Lisboa e Porto perdem parte da hegemonia neste índice a favor das respectivas cinturas concelhias. Prova disso, o município de Lisboa mantém-se no topo da distribuição, mas a uma curta distância dos de Sintra e de Loures. O município do Porto remete-se para uma posição secundária face à de concelhos como Vila Nova de Gaia, Guimarães, Braga, Matosinhos, Gondomar, Paredes ou Vila Nova de Famalicão, entre outros. Esta situação denota que a valores elevados de densidade (populacional e habitacional) nem sempre correspondem valores equivalentes de crescimento ou vitalidade demográfica.

Abaixo das posições hegemónicas referidas, há alguns concelhos das NUTS III Baixo Vouga, Baixo Mondego e Pinhal Litoral, especialmente Leiria e Coimbra, que conseguem sobrepor-se à extensa faixa interior que, de Alto Trás-os-Montes ao sul do Alentejo, forma uma grande área de debilidade demográfica generalizada. Na faixa concelhia mais interior não há nenhum concelho (sedes de distrito incluídas) com valor acima da média continental, e só Viseu e Vila Real conseguem aproximar-se dos índices mais elevados que caracterizam as cinturas metropolitanas. A conjugação generalizada de fracas densidades populacional e habitacional e de baixas taxas de crescimento e de vitalidade demográfica no interior do Continente explica aquele resultado. No Algarve o retrato global é algo mais satisfatório do que o da generalidade concelhia do interior, mas também não há nenhum município que se destaque de forma significativa: Albufeira e Faro são os concelhos com melhores *scores*, mas cerca de 9% abaixo da média do Continente.

No índice de saúde e assistência social incorporam-se indicadores relacionados com o pessoal ao serviço, equipamentos, infra-estruturas e acção desenvolvida em estabelecimentos de saúde e de segurança social. Este índice é o que apresenta a média não ponderada dos valores concelhios mais baixa, o que se justifica por uma grande proliferação de *scores* baixos. Simultaneamente, é o índice que detém o mais alto coeficiente de variação e a maior amplitude nas distribuições dos valores regionais e concelhios. Estas características respeitam a grande concentração de valores

consideravelmente elevados em poucos concelhos e a maior assimetria generalizada verificadas nos indicadores originais que compõem este índice. Entre todos os índices calculados, o da saúde e assistência social é o que apresenta o menor número de concelhos (27) acima da média continental e é, também, o que mais distancia os três concelhos melhor posicionados (Porto, Lisboa e Coimbra) dos restantes. Esta situação explica-se, em parte, pelo tipo e pela localização de alguns dos equipamentos e serviços que os indicadores originais deste índice medem (designadamente, dos equipamentos hospitalares). Não obstante, é a existência e a localização destes equipamentos que produzem parte substancial das diferenças existentes entre concelhos e regiões e não podem, pois, deixar de ser considerados.

Ao nível das NUTS II, Lisboa e Vale do Tejo persiste com o maior valor (124,6), em claro contraste com o Algarve, que assume, neste caso, o resultado mais desfavorável (74,6). Independentemente da grande assimetria que retrata, o índice de saúde e assistência social é dos que menos diferenças regista entre as zonas litorais e interiores do Continente. Para além da hegemonia das NUTS III da Grande Lisboa, Baixo Mondego e Grande Porto, as únicas com valores acima da média do Continente, constata-se neste índice um papel importante assumido por parte considerável do interior continental. Beira Interior Norte, Cova da Beira e Beira Interior Sul, com resultados superiores às de todas as outras NUTS III não hegemónicas, são os exemplos mais notórios.

Simultaneamente, a análise concelhia do índice de saúde e assistência social permite constatar que dos 27 municípios situados acima da média continental, 16 são sedes de distrito, metade dos quais localizados no interior. Se a este resultado se adicionar toda a mancha concelhia vizinha dos principais municípios do interior, encontra-se uma extensa faixa territorial interior, de norte a sul, visivelmente mais dotada do que algumas NUTS III e concelhos do litoral. Certamente que a capacidade de absorção do investimento público – e não só – das capitais de distrito e os dividendos para as zonas do interior daí decorrentes ajudam a explicar aqueles resultados. Seja como for, as áreas geográficas menos dotadas no sector da saúde e assistência social não se encontram especialmente concentradas nas faixas mais interiores. Pelo contrário, proliferam de norte a sul do território continental, de uma forma algo difusa. Salvaguardando os casos mais favoráveis dos concelhos sede de distrito, ressalta deste índice a percepção das carências generalizadas do território continental no âmbito da saúde e assistência social.

O índice educação e cultura é calculado a partir de informação sobre estabelecimentos de ensino, incluindo número de escolas, professores e alunos matriculados, bibliotecas, espectáculos públicos e despesas municipais com actividades culturais. A distribuição concelhia mostra-se bastante correlacionada positivamente com a do índice de saúde e assistência social, o que denota características comuns entre ambas. Apesar disso, este índice apresenta uma média não ponderada dos valores concelhios mais elevada do que o anterior e é bastante menos assimétrico. Significa isto que, no sector da educação e cultura, os concelhos apresentam valores globalmente mais elevados e menos díspares entre si do que no domínio da saúde e assistência social.

Lisboa e Vale do Tejo e Algarve são na educação e cultura, tal como no índice anterior, as NUTS II com os valores mais elevado (109,0) e mais baixo (92,3),

respectivamente. Veja-se, no entanto, como a amplitude entre estes dois valores extremos é mais reduzida do que no índice anterior, sendo acompanhada de uma maior aproximação entre as posições intermédias das restantes NUTS II. Nos valores desagregados, as situações de destaque da Grande Lisboa, Baixo Mondego e Grande Porto são acrescidas, neste caso, pelas do Alentejo Central e do Cávado, todas acima da média continental. Um segundo patamar de desenvolvimento é formado por uma faixa interior que se estende de Dão-Lafões e Beira Interior Norte até ao Algarve, com ramificação na Península de Setúbal.

Por concelhos constata-se que na educação e cultura Lisboa, Coimbra e Porto encontram-se nas três primeiras posições, mas nota-se, igualmente, alguma preponderância generalizada dos municípios sede de distrito – que não é alheia, certamente, à localização dos estabelecimentos universitários, entre outros de maior dimensão. Entre os 25 primeiros concelhos da distribuição definida por este índice 13 são capitais de distrito e, destes, Évora, Beja, Bragança, Vila Real, Viseu e Guarda estão claramente acima da média. Destes e de outros casos similares, ressalta uma proximidade relativa dos valores entre as zonas litorais e as zonas interiores do Continente e uma diferenciação espacial de dotações na educação e cultura pouco consentânea com uma divisão vertical do território. O retrato concelhio global do Continente apresenta, neste domínio, uma maior concentração de baixos valores a norte do rio Tejo e, de uma forma particular, nas zonas delimitadas pelas NUTS III do Pinhal Interior Sul, Tâmega, Entre Douro e Vouga e Oeste.

O índice de rendimento tem origem em indicadores referentes ao rendimento das populações, incluindo indivíduos economicamente activos, trabalhadores por conta de outrem e pensionistas, mercado monetário associado a particulares e parque automóvel. A óptica dominante desta medida é o rendimento e não o consumo. Onde, no plano metodológico, os *clusters* regionais e concelhios que sobressaem deste índice são relativamente independentes dos principais mercados do Continente, ainda que os possam reflectir.

Entre as NUTS II o nível médio de maior rendimento localiza-se em Lisboa e Vale do Tejo (120,5), caso único acima da média continental. Por oposição, o Alentejo é a NUTS II com o rendimento médio mais baixo (85,6), não muito longe dos do Norte (88,6) e do Centro (89,2). O Algarve consegue neste domínio um dos seus melhores *scores* (95,8), embora notoriamente abaixo do valor de Lisboa e Vale do Tejo.

No plano das NUTS III, os níveis médios de rendimento na Grande Lisboa, Península de Setúbal, Grande Porto e Baixo Mondego superam os de todas as outras, sendo as únicas com valores acima da média 100. A Grande Lisboa, com os municípios de Lisboa e Oeiras à cabeça, detém neste índice um dos seus valores mais elevados. Grosso modo, esta hegemonia estende-se a grande parte da faixa concelhia continental, exceptuando-se alguns municípios localizados predominantemente nas NUTS III do Minho-Lima, Cávado, Oeste e Alentejo Litoral. No interior encontram-se níveis de rendimento mais elevados nos concelhos sede de distrito, com especial destaque para o diferencial existente no Alentejo. Évora, Portalegre e Beja são os concelhos do interior continental cujos valores, neste índice, mais se aproximam dos valores dos concelhos com maior rendimento no litoral. Não obstante, o facto deste índice estar menos associado ao investimento público (comparativamente aos da saúde

e da educação), faz com que os concelhos que integram as capitais de distrito – e o interior em geral – não sobressaiam tanto como nos dois índices anteriores. Reflexo disso, os níveis médios de rendimento no interior são notoriamente deficitários face à hegemonia de grande parte da faixa litoral, com o Algarve incluído. As maiores bolsas concelhias de rendimento mais baixo localizam-se, predominantemente, numa extensa área central que ocupa parte considerável do Norte e do Centro e no sudeste alentejano.

O índice de emprego e actividade económica constitui-se originariamente por indicadores de volume e qualificação de emprego, desemprego, empresas sediadas e constituídas e actividade económica. Este índice e o de rendimento apresentam o maior coeficiente de correlação existente entre os índices parciais, o que lhes descobre importantes características comuns e, certamente, alguma relação causal.

Lisboa e Vale do Tejo (116,6) e Algarve (107,7) são as duas únicas NUTS II com valores acima da média continental no emprego e actividade económica. Por contraste, o Alentejo denota a situação genericamente mais desfavorável, com um índice de 81,5. O Centro e o Norte, com índices de 92,6 e 89,9, respectivamente, estão próximos entre si e revelam neste domínio um desempenho semelhante ao constatado no índice anterior.

Desagregados os valores, os índices mais elevados encontram-se em quase toda a costa atlântica, com particular excepção de alguns concelhos do Minho-Lima e do Alentejo Litoral. O Algarve assume neste domínio uma importância substancial, com 9 dos seus 16 concelhos acima da média continental. Este índice é, aliás, o que retrata uma maior hegemonia das regiões e concelhos litorais. A faixa atlântica que permite desenhar é mais densa, mais estreita e mais alongada do que a dos outros índices. No interior as manchas de menor índice de emprego e actividade económica localizam-se, maioritariamente, nas zonas centrais das NUTS II do Norte e Centro, no nordeste e sudeste alentejanos.

---

### 3.2. DESENVOLVIMENTO E REGIÕES

---

Entre as Regiões, Lisboa e Vale do Tejo é, de acordo com os dados apurados, a NUTS II mais desenvolvida do Continente, assumindo os maiores valores no índice global de desenvolvimento e nos cinco índices parciais. O seu melhor desempenho exerce-o no sector da saúde e assistência social, ao passo que o resultado menos favorável é o da educação e cultura. Internamente, a Grande Lisboa distingue-se claramente das restantes NUTS III em todos os índices. Além desta, a Península de Setúbal assume também um papel importante, designadamente nos sectores do rendimento e da demografia. À escala municipal, o índice global de desenvolvimento coloca um grupo de 10 concelhos com valores iguais ou acima da média continental: Lisboa, Oeiras, Cascais, Amadora, Almada, Sintra, Setúbal, Barreiro, Entroncamento e Loures.

O Norte apresenta o segundo valor das NUTS II mais elevado no índice global de desenvolvimento, mas já abaixo do limiar 100. Consegue superar este limiar, ainda que ligeiramente, no índice da demografia, onde conta com 5 NUTS III

consideravelmente fortes – Grande Porto, Ave, Cávado, Entre Douro e Vouga e Tâmega. A saúde e assistência social é a sua área mais deficitária. O Grande Porto lidera em quase todas as áreas, com excepção no emprego e actividade económica onde a NUTS III Entre Douro e Vouga apresenta um valor ligeiramente superior. Entre os concelhos há 6 que superam a média 100 no índice global: Porto, São João da Madeira, Braga, Vila Nova de Gaia, Maia e Matosinhos.

O Algarve ocupa a terceira posição no índice global de desenvolvimento, está acima da média continental na área do emprego e actividade económica e, tal como o Norte, tem o mais baixo valor regional no índice da saúde e assistência social, aliás, o mais baixo de entre as NUTS II. Ao nível concelhio (NUTS II e NUTS III coincidem nos limites geográficos), Faro apresenta o nível de maior desenvolvimento global, tem um *score* muito elevado na saúde e assistência social, mas perde a liderança para Albufeira na demografia e no emprego e actividade económica. Além destes dois concelhos, Portimão, Lagos e Loulé assumem, também, alguma preponderância na generalidade dos índices.

Segue-se a NUTS II do Centro, com o quarto melhor resultado global, sem conseguir atingir a média continental em nenhum dos índices parciais. O seu melhor desempenho encontra-se na educação e cultura e é mais deficitário na demografia. A NUTS III mais desenvolvida é a do Baixo Mondego, designadamente na área da saúde e assistência social, mas perde a liderança para a do Pinhal Litoral no emprego e actividade económica e na demografia. Ao nível concelhio registam-se 5 casos acima da média continental no índice global: Coimbra, Aveiro, Viseu, Guarda e Leiria – todos sede de distrito.

O Alentejo regista, entre as NUTS II, o índice global mais baixo do Continente, fruto, em grande medida, da debilidade demográfica que caracteriza esta região. Logra o seu melhor resultado na área da educação e cultura, o segundo maior do Continente. Entre as NUTS III, o Alentejo Central é a que assume o índice global mais elevado, mas na saúde e assistência social cede a primeira posição ao Alto Alentejo e na demografia ao Alentejo Litoral. Nos municípios há apenas dois – Évora e Beja – que superam a média 100 no índice global, com Portalegre e Sines perto desse desiderato.

## REFERÊNCIAS BIBLIOGRÁFICAS

BOUROCHE, JEAN-MARIE; SAPORTA, GILBERT (1989), *L'Analyse des Données*, Press Universitaires de France, Paris, 4<sup>a</sup> ed.

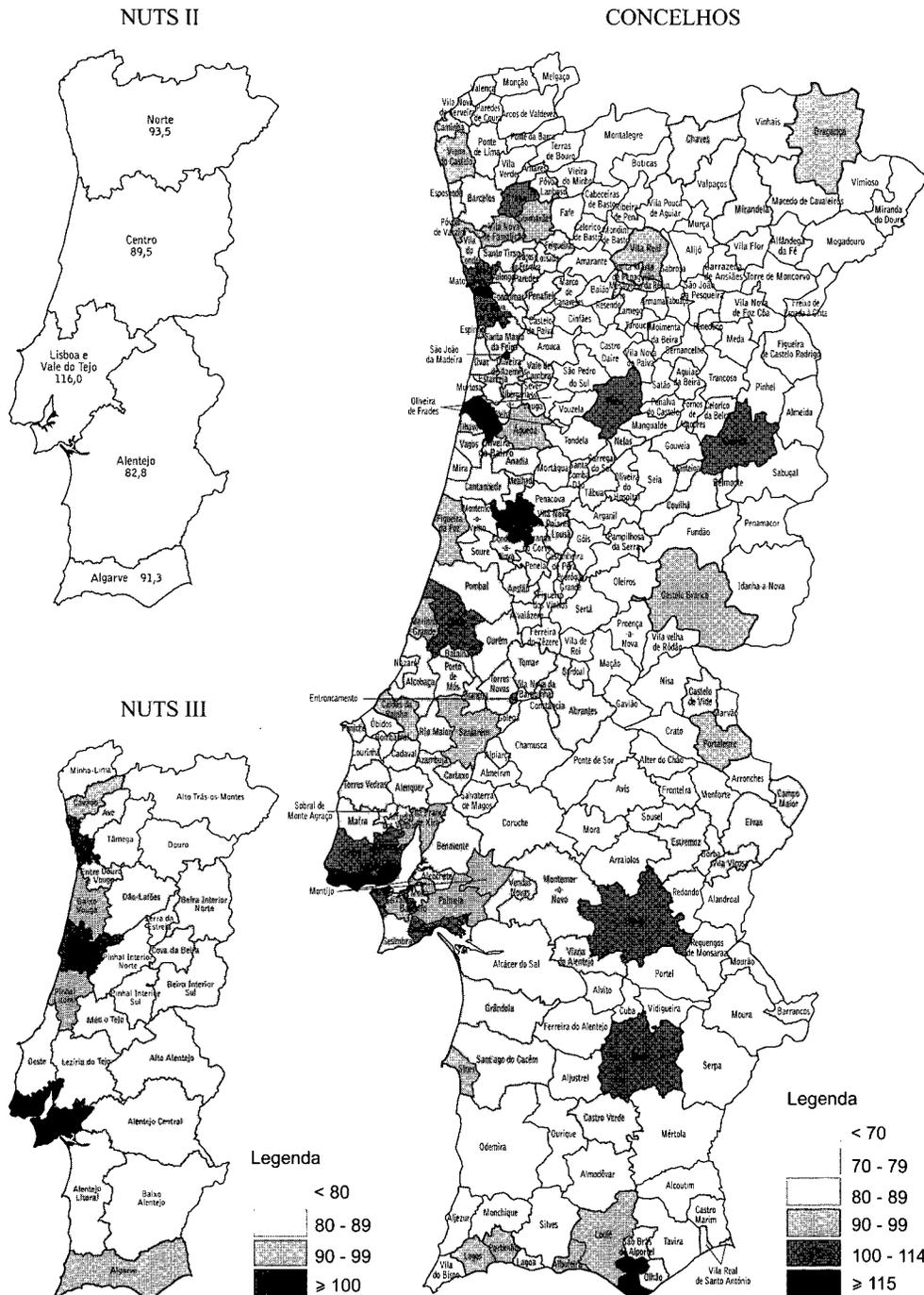
BRANDÃO, ANA; PIRES, ANA; PORTUGAL, JORGE (1998), “Agrupamentos de Concelhos de Portugal Continental e Sua Caracterização”, *Revista de Estatística*, INE, Lisboa, 1º quadrimestre, pp. 71-93

DILLON, WILLIAM R.; GOLDSTEIN, MATTHEW (1984), *Multivariate Analysis. Methods and Applications*, John Wiley & Sons, New York

- HAIR, JOSEPH F., *et al* (1992), *Multivariate Data Analysis*, Macmillan Publishing Company, New York
- INE (1993), *Estudo Sobre o Poder de Compra Concelhio*, INE/Direcção Regional do Centro, Coimbra
- INE (1999), *Nomenclaturas Territoriais 1998. Designações e Códigos*, Lisboa
- INE (2000), *Indicadores de Desenvolvimento Económico e Social das Comarcas / Concelhos da Galiza e do Norte de Portugal*, INE/Direcção Regional do Norte, Porto
- KINNEAR, PAUL R.; GRAY, COLIN D. (1994), *SPSS for Windows. Made Simple*, Erlbaum Taylor & Francis, Hove
- NORUSIS, MARIJA J.; SPSS INC. (1994), *SPSS Professional Statistics 6.1*, SPSS INC., Chicago
- PESTANA, MARIA HELENA; GAGEIRO, JOÃO NUNES (2000), *Análise de Dados para as Ciências Sociais. A Complementaridade do SPSS*, Edições Sílabo, Lisboa, 2ª ed.
- REIS, ELISABETH (1997), *Estatística Multivariada Aplicada*, Edições Sílabo, Lisboa

ANEXOS

ÍNDICE DE DESENVOLVIMENTO  
(Portugal Continental = 100)



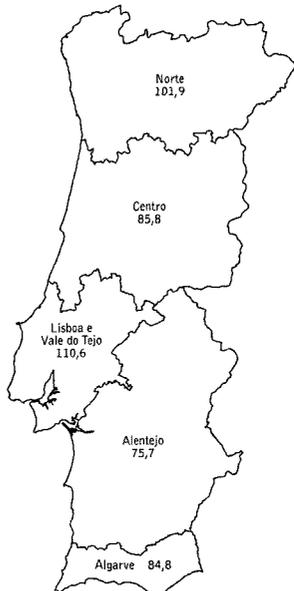
Índice de Desenvolvimento					
Lisboa	206,3	Castelo Branco	91,5	Vila Nova de Poiares	83,6
Porto	165,4	<b>ALGARVE</b>	91,3	BEIRA INTERIOR SUL	83,6
Coimbra	142,7	Guimarães	91,3	Anadia	83,5
GRANDE LISBOA	136,4	BAIXO VOUGA	91,2	Felgueiras	83,3
São João da Madeira	121,1	Viana do Castelo	90,8	Cantanhede	83,3
Oeiras	118,6	Montijo	90,5	Albergaria-a-Velha	82,8
Faro	117,8	<b>CENTRO</b>	89,5	Arruda dos Vinhos	82,8
<b>LISBOA E VALE DO TEJO</b>	116,0	ENTRE DOURO E VOUGA	89,1	<b>ALENTEJO</b>	82,8
Cascais	115,9	Vila Real de Santo António	89,0	COVA DA BEIRA	82,7
Aveiro	115,5	Alcochete	88,9	Cartaxo	82,6
GRANDE PORTO	113,7	Sesimbra	88,8	LEZÍRIA DO TEJO	82,4
Évora	113,4	Mafra	88,8	Penafiel	82,3
Braga	111,9	Santa Maria da Feira	88,7	Rio Maior	82,3
Amadora	110,7	Torres Vedras	88,7	Paredes	82,2
Vila Nova de Gaia	109,8	Valongo	88,6	Moita	82,1
Almada	109,2	Oliveira de Azeméis	88,4	Mirandela	82,1
BAIXO MONDEGO	108,5	ALENTEJO CENTRAL	88,3	Vagos	82,1
Sintra	107,6	Elvas	88,0	Caminha	81,9
Maia	106,4	Covilhã	87,8	Abrantes	81,9
Setúbal	106,1	São Brás de Alportel	87,6	BEIRA INTERIOR NORTE	81,8
Barreiro	104,3	Ovar	87,5	Campo Maior	81,5
Beja	103,3	Ílhavo	87,3	Nelas	81,4
Viseu	102,5	AVE	87,2	ALTO ALENTEJO	81,4
Matosinhos	102,4	Lagoa	86,8	DÃO-LAFÕES	81,3
Entroncamento	102,2	Vila do Conde	86,6	Peniche	81,2
Guarda	101,6	Tomar	86,6	Vale de Cambra	81,1
Leiria	101,3	Vila Viçosa	86,4	Olhão	80,8
<b>PORTUGAL CONTINENTAL</b>	100,0	Gondomar	86,2	Peso da Régua	80,6
Loures	99,6	Barcelos	86,0	Estremoz	80,5
<b>PENÍNSULA DE SETÚBAL</b>	99,5	Estarreja	85,9	Alenquer	80,4
Vila Franca de Xira	99,1	Condeixa-a-Nova	85,8	Montemor-o-Novo	80,4
Bragança	98,7	Vendas Novas	85,7	Porto de Mós	80,3
Portimão	98,5	Alcanena	85,5	Santiago do Cacém	80,3
Vila Real	98,1	Santo Tirso	85,3	ALENTEJO LITORAL	80,3
Seixal	97,8	Alcobaça	85,3	Ansião	80,1
Albufeira	97,7	Batalha	85,1	Castelo de Vide	79,8
Portalegre	97,3	Ourém	85,0	Viana do Alentejo	79,8
Sines	96,9	Sobral de Monte Agraço	84,9	Constância	79,8
Póvoa de Varzim	96,4	Benavente	84,9	Lourinhã	79,0
Lagos	95,8	Grândola	84,8	Miranda do Corvo	78,9
Loulé	94,7	Lousã	84,8	BAIXO ALENTEJO	78,9
Figueira da Foz	94,4	Oliveira do Bairro	84,7	Lousada	78,8
Santarém	94,0	OESTE	84,6	Alcácer do Sal	78,8
Marinha Grande	93,7	Paços de Ferreira	84,5	Fafe	78,4
Caldas da Rainha	93,5	Castro Verde	84,5	MINHO-LIMA	78,4
<b>NORTE</b>	93,5	Pombal	84,3	Nazaré	78,3
CÁVADO	93,5	Azambuja	84,3	Macedo de Cavaleiros	77,9
PINHAL LITORAL	93,2	Torres Novas	84,2	Lamego	77,9
Águeda	92,4	Esposende	83,9	TÂMEGA	77,3
Espinho	92,1	Chaves	83,9	ALTO TRÁS-OS-MONTES	77,3
Palmela	92,1	MÉDIO TEJO	83,8	Pampilhosa da Serra	77,1
Vila Nova de Famalicão	91,7	Mealhada	83,8	Miranda do Douro	76,9

## Índice de Desenvolvimento (continuação)

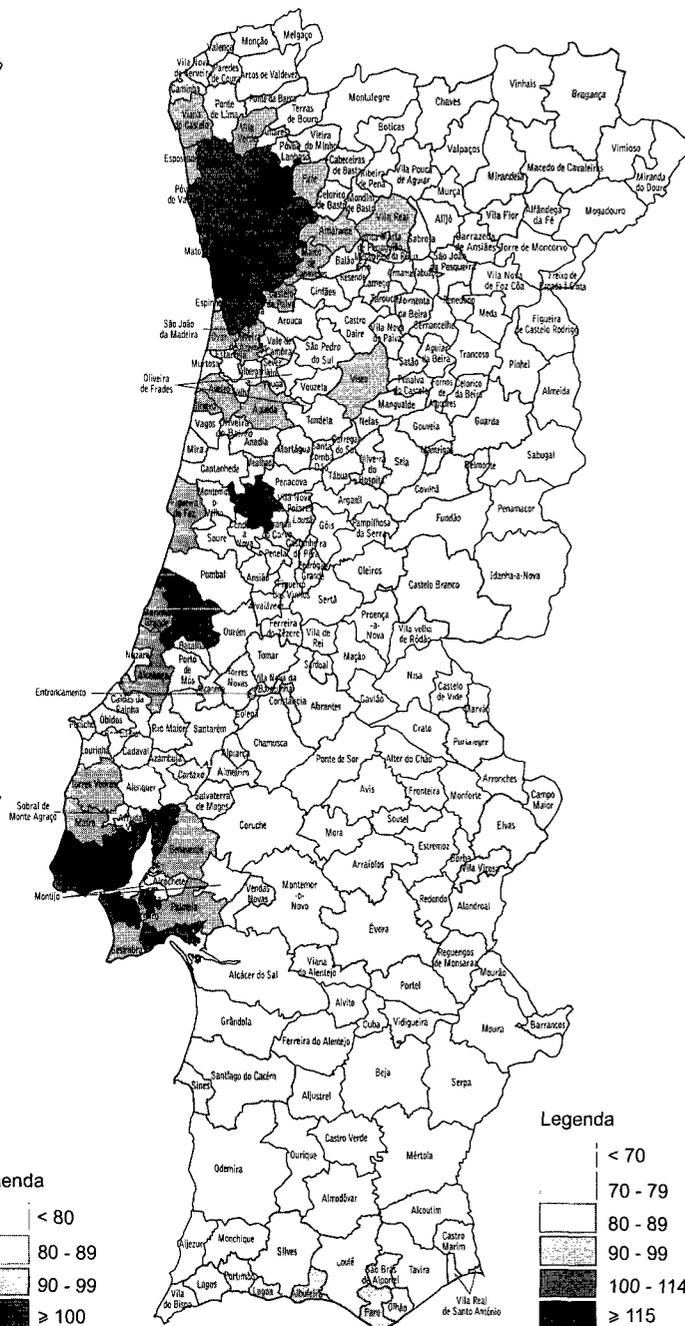
Seia	76,9	Avis	72,8	Alijó	68,3
Alvito	76,9	Penedono	72,5	Vila Pouca de Aguiar	68,2
Manteigas	76,9	Vila Nova de Foz Côa	72,4	Melgaço	68,2
Almeirim	76,7	Mogadouro	72,3	São Pedro do Sul	67,9
Amarante	76,7	Arraiolos	72,1	Proença-a-Nova	67,7
Oliveira do Hospital	76,5	Coruche	72,1	Ourique	67,5
Trancoso	76,4	Monchique	71,9	Ferreira do Alentejo	67,5
DOURO	76,4	Figueiró dos Vinhos	71,9	Murça	67,5
Mangualde	76,4	Alpiarça	71,7	Sátão	67,3
Monforte	76,2	Figueira de Castelo Rodrigo	71,7	Vidigueira	67,3
Gouveia	76,1	Aguiar da Beira	71,7	Carraceda de Ansiães	67,0
Marco de Canavezes	76,1	Ponte da Barca	71,7	Castro Daire	66,8
Silves	76,0	Moura	71,5	PINHAL INTERIOR SUL	66,7
Penacova	76,0	Vila Verde	71,3	Arronches	66,5
Fundão	75,9	Alfândega da Fé	71,1	Penalva do Castelo	66,4
PINHAL INTERIOR NORTE	75,8	Vila do Bispo	71,0	Marvão	66,3
Reguengos de Monsaraz	75,8	Mesão Frio	70,9	Ferreira do Zêzere	66,3
Mortágua	75,8	Pinhel	70,6	Sabrosa	66,3
SERRA DA ESTRELA	75,7	Vieira do Minho	70,5	Terras de Bouro	66,2
Arouca	75,6	Vila Velha de Ródão	70,3	Meda	66,2
Ponte de Sôr	75,4	Serpa	70,2	Sabugal	66,0
Fronteira	75,3	Cuba	70,2	Vila de Rei	66,0
Belmonte	75,2	Mora	70,1	Ribeira de Pena	66,0
Golegã	75,2	Castro Marim	70,1	Armamar	66,0
Murtosa	75,1	Sousel	70,0	Arcos de Valdevez	65,9
Arganil	75,0	Almodôvar	69,9	Vimioso	65,9
Soure	74,7	Sertã	69,9	Mondim de Basto	65,7
Oliveira de Frades	74,7	Odemira	69,8	Resende	65,5
Castelo de Paiva	74,6	Aljezur	69,8	Barrancos	65,5
Valença	74,5	Tábua	69,8	Celorico de Basto	65,4
Bombarral	74,5	Crato	69,8	Mértola	65,3
Mira	74,4	Alvaiázere	69,7	Portel	64,7
Vila Nova da Barquinha	74,4	Vila Flor	69,7	Cinfães	64,6
Montemor-o-Velho	74,4	Carregal do Sal	69,6	Santa Marta de Penaguião	64,5
Almeida	74,3	Nisa	69,6	Freixo de Espada à Cinta	64,3
Santa Comba Dão	74,2	Castanheira de Pera	69,3	Idanha-a-Nova	64,2
Sever do Vouga	74,2	Sardoal	69,3	Alcoutim	63,7
Alter do Chão	74,2	Vila Nova de Paiva	69,3	Valpaços	63,6
Ponte de Lima	74,2	Redondo	69,1	Mação	63,5
Mourão	74,1	Montalegre	69,1	Sernancelhe	63,4
Tavira	74,0	Góis	68,9	Tabuaço	63,1
Aljustrel	73,7	Penela	68,8	Baião	62,9
Tondela	73,7	Óbidos	68,7	Penamacor	62,5
Cadaval	73,6	Tarouca	68,7	Boticas	61,9
Salvaterra de Magos	73,5	Vouzela	68,7	Alandroal	61,6
Torre de Moncorvo	73,4	Cabeceiras de Basto	68,7	Vinhais	61,4
Pedrógão Grande	73,2	São João da Pesqueira	68,6	Oleiros	61,3
Celorico da Beira	73,1	Moimenta da Beira	68,5	Gavião	61,2
Amares	73,1	Paredes de Coura	68,4		
Borba	73,0	Fornos de Algodres	68,4		
Vila Nova de Cerveira	72,9	Chamusca	68,4		
Póvoa de Lanhoso	72,9	Monção	68,4		

## ÍNDICE DE DEMOGRAFIA (Portugal Continental = 100)

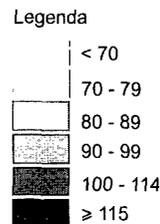
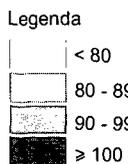
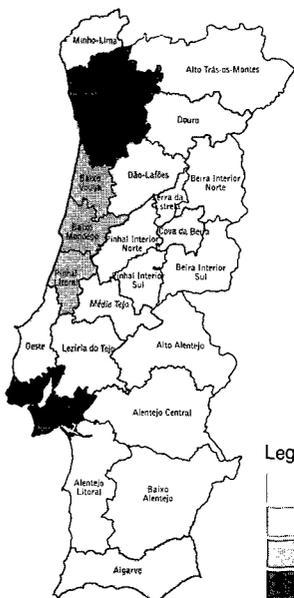
NUTS II



CONCELHOS



NUTS III



Índice de Demografia					
Lisboa	134,8	Espinho	94,9	DÃO-LAFÕES	84,9
Sintra	133,8	Viana do Castelo	94,6	Cartaxo	84,8
Loures	131,4	PINHAL LITORAL	94,6	ALGARVE	84,8
GRANDE LISBOA	125,1	Palmela	94,4	Nazaré	84,7
Vila Nova de Gaia	121,9	Vila Verde	93,9	Chaves	84,6
Seixal	119,7	São João da Madeira	93,8	Mira	84,5
Guimarães	114,7	Marinha Grande	93,0	Celorico de Basto	84,4
Braga	114,3	Entroncamento	92,7	Mondim de Basto	84,3
Matosinhos	112,0	Águeda	92,4	Mangualde	84,2
Gondomar	111,2	Torres Vedras	92,1	Lousã	84,2
GRANDE PORTO	110,9	Castelo de Paiva	91,5	Anadia	83,9
LISBOA E VALE DO TEJO	110,6	Sesimbra	91,5	Lagos	83,8
Paredes	110,5	Albufeira	91,4	Lourinhã	83,8
Vila Nova de Famalicão	109,8	BAIXO VOUGA	91,3	Guarda	83,8
Amadora	109,7	BAIXO MONDEGO	91,2	Vila Real de Santo António	83,6
Vila Franca de Xira	109,3	Vila Real	91,0	Alenquer	83,5
Oeiras	109,1	Mafra	90,8	Bragança	83,4
Santa Maria da Feira	108,8	Alcobaça	90,6	Salvaterra de Magos	83,3
Valongo	108,8	Faro	90,6	Baião	83,2
Cascais	108,7	Benavente	90,5	Santarém	83,1
Barcelos	108,4	Figueira da Foz	90,4	Mealhada	83,1
Maia	108,4	Sines	89,4	Santiago do Cacém	83,0
Paços de Ferreira	107,9	Portimão	89,3	Castelo Branco	82,9
Felgueiras	107,3	Póvoa de Lanhoso	89,3	Valença	82,9
Porto	107,3	Peso da Régua	88,9	Oliveira do Bairro	82,8
AVE	107,2	Lagoa	88,7	Cabeceiras de Basto	82,7
Lousada	107,0	Caldas da Rainha	88,6	DOURO	82,7
CÁVADO	106,6	Ponte de Lima	88,6	Arruda dos Vinhos	82,5
Santo Tirso	105,1	Amares	88,6	Rio Maior	82,5
Penafiel	104,9	Peniche	88,1	LEZÍRIA DO TEJO	82,4
Almada	104,5	Vagos	87,9	Macedo de Cavaleiros	82,4
PENÍNSULA DE SETÚBAL	102,8	Tarouca	87,8	Sobral de Monte Agraço	82,3
NORTE	101,9	Loulé	87,8	Azambuja	82,3
Leiria	101,9	Albergaria-a-Velha	87,6	Santa Marta de Penaguião	82,1
Póvoa de Varzim	101,9	Estarreja	87,6	Mesão Frio	82,1
Vila do Conde	101,4	Porto de Mós	87,5	Caminha	81,9
ENTRE DOURO E VOUGA	100,7	Vale de Cambra	87,4	Constância	81,8
Coimbra	100,7	OESTE	87,4	Nelas	81,6
TÂMEGA	100,1	Montijo	86,8	Cinfães	81,5
PORTUGAL CONTINENTAL	100,0	Lamego	86,7	Sátão	81,4
Moita	100,0	Évora	86,6	MÉDIO TEJO	81,2
Setúbal	99,8	Pombal	86,3	Montemor-o-Velho	81,2
Ovar	99,3	Olhão	86,2	Tomar	81,2
Marco de Canavezes	99,2	Ourém	86,1	Torres Novas	81,2
Esposende	99,2	CENTRO	85,8	COVA DA BEIRA	81,0
Amarante	99,0	Cantanhede	85,6	Alcanena	81,0
Viseu	98,0	Batalha	85,5	Moimenta da Beira	81,0
Fafe	97,6	Arouca	85,4	Mirandela	80,9
Oliveira de Azeméis	97,6	Covilhã	85,3	Vila Pouca de Aguiar	80,8
Ílhavo	97,3	Alcochete	85,3	Miranda do Corvo	80,8
Aveiro	96,8	Almeirim	85,0	Tondela	80,8
Barreiro	96,6	MINHO-LIMA	85,0	Óbidos	80,8

<b>Índice de Demografia (continuação)</b>					
Sever do Vouga	80,5	Borba	75,7	Vila Nova de Foz Côa	69,1
Seia	80,5	Castro Daire	75,5	Castanheira de Pera	69,0
Vieira do Minho	80,4	Fundão	75,5	Fornos de Algodres	69,0
São João da Pesqueira	80,3	Murça	75,3	Fronteira	69,0
Ponte da Barca	80,3	Cadaval	75,3	Meda	68,9
Vila Viçosa	80,2	Alcácer do Sal	75,2	Freixo de Espada à Cinta	68,5
Elvas	80,0	Moura	75,1	Almeida	68,3
Vendas Novas	79,9	BEIRA INTERIOR SUL	75,1	Trancoso	68,1
Vila Flor	79,9	São Brás de Alportel	75,1	Penela	67,8
Abrantes	79,8	Soure	75,0	Avis	67,7
Aljô	79,8	Alpiarça	74,8	Sousel	67,6
Oliveira do Hospital	79,8	Ansião	74,8	Proença-a-Nova	67,6
Silves	79,6	Belmonte	74,6	Viana do Alentejo	67,6
Mortágua	79,6	Penedono	74,3	Torre de Moncorvo	67,5
Portalegre	79,5	Arcos de Valdevez	74,3	Figueira de Castelo Rodrigo	67,3
Penacova	79,4	Bombarral	74,2	Ferreira do Zêzere	67,2
Tabuaço	79,2	Ponte de Sôr	74,2	Alandroal	67,2
Vila Nova da Barquinha	79,1	Tábua	74,1	Vidigueira	67,0
Santa Comba Dão	79,1	Boticas	73,9	Melgaço	67,0
Armamar	79,0	BAIXO ALENTEJO	73,8	Vimioso	66,9
Resende	78,9	PINHAL INTERIOR NORTE	73,8	Mora	66,3
ALTO TRÁS-OS-MONTES	78,8	Mogadouro	73,7	Aljezur	66,1
Beja	78,8	Penalva do Castelo	73,6	PINHAL INTERIOR SUL	66,1
Condeixa-a-Nova	78,8	Mourão	73,6	Castelo de Vide	66,0
ALENTEJO LITORAL	78,5	Alfândega da Fé	73,5	Monchique	65,6
São Pedro do Sul	78,4	BEIRA INTERIOR NORTE	73,2	Monforte	65,6
Vila do Bispo	78,3	Miranda do Douro	73,0	Ourique	65,5
Campo Maior	78,3	Ferreira do Alentejo	73,0	Sardoal	65,4
Sernancelhe	78,2	ALTO ALENTEJO	72,9	Alvaiázere	64,9
Aljustrel	78,1	Carraceda de Ansiães	72,8	Alvito	64,6
Oliveira de Frades	78,1	Odemira	72,7	Mértola	64,3
Castro Verde	77,8	Redondo	72,7	Alter do Chão	63,9
Terras de Bouro	77,8	Paredes de Coura	72,5	Crato	63,4
Ribeira de Pena	77,7	Almodôvar	72,5	Góis	63,1
Carregal do Sal	77,7	Montemor-o-Novo	72,4	Mação	62,0
Murtosa	77,7	Chamusca	72,2	Pedrogão Grande	61,5
ALENTEJO CENTRAL	77,6	Portel	72,2	Nisa	61,3
Valpaços	77,4	Celorico da Beira	72,0	Marvão	61,3
Vila Nova de Paiva	77,2	Gouveia	71,5	Alcoutim	61,2
Vila Nova de Cerveira	76,9	Estremoz	71,5	Oleiros	60,6
Vila Nova de Poiares	76,7	Serpa	71,5	Arronches	60,3
Aguiar da Beira	76,6	Cuba	71,2	Sabugal	59,9
Tavira	76,6	Figueiró dos Vinhos	71,0	Vila Velha de Ródão	58,7
Coruche	76,4	Sertã	70,9	Vila de Rei	58,2
SERRA DA ESTRELA	76,3	Castro Marim	70,6	Gavião	57,9
Vouzela	76,3	Grândola	70,3	Idanha-a-Nova	57,6
Reguengos de Monsaraz	76,3	Barrancos	70,1	Penamacor	54,3
Golegã	76,1	Montalegre	69,8	Pampilhosa da Serra	54,2
Monção	75,9	Arganil	69,8		
Sabrosa	75,9	Vinhais	69,7		
Manteigas	75,8	Pinhel	69,7		
ALENTEJO	75,7	Arraiolos	69,6		



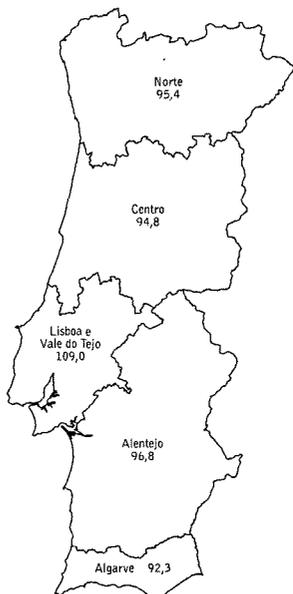
<b>Índice de Saúde e Assistência Social</b>					
Porto	394,9	Oliveira do Bairro	84,9	São Brás de Alportel	67,0
Lisboa	384,7	Matosinhos	84,8	Soure	66,7
Coimbra	280,9	Alcácer do Sal	84,7	Vila Nova de Foz Côa	66,6
Faro	177,7	Torres Vedras	84,6	Fronteira	66,3
GRANDE LISBOA	167,3	ALENTEJO CENTRAL	84,5	Nisa	66,3
BAIXO MONDEGO	164,6	Lagos	84,0	Castanheira de Pera	66,1
Portalegre	159,8	Góis	83,4	AVE	65,8
Beja	156,5	Alfândega da Fé	82,9	PINHAL LITORAL	65,5
Guarda	155,2	Montijo	82,7	Sabugal	65,3
Amadora	155,2	Entroncamento	82,1	Celorico da Beira	65,2
Évora	140,6	PENÍNSULA DE SETÚBAL	81,1	DÃO-LAFÕES	64,8
GRANDE PORTO	134,6	CÁVADO	81,0	Vila Velha de Ródão	64,7
Aveiro	132,2	Alvito	80,9	Paredes de Coura	64,6
São João da Madeira	131,0	Torres Novas	80,4	LEZÍRIA DO TEJO	64,5
Bragança	130,4	Peso da Régua	79,9	Ovar	64,2
Barreiro	127,3	Montemor-o-Novo	79,7	Belmonte	64,2
Vila Real	126,0	Guimarães	79,6	Alcobaça	64,1
<b>LISBOA E VALE DO TEJO</b>	124,6	Alcoutim	79,4	Montemor-o-Velho	63,8
Monforte	122,9	Arganil	79,4	Vila Flor	62,7
Viseu	122,9	Azambuja	79,3	OESTE	62,6
Pampilhosa da Serra	116,3	Lamego	79,3	Arraiolos	62,5
Castelo Branco	116,0	<b>ALENTEJO</b>	78,7	Vimioso	62,4
Santarém	113,7	SERRA DA ESTRELA	78,6	Vila Viçosa	61,9
Setúbal	112,6	Scia	78,2	Sines	61,8
Braga	112,3	Águeda	78,0	ALENTEJO LITORAL	61,8
Condeixa-a-Nova	107,8	Fundão	77,7	Constância	61,4
Covilhã	106,7	Vila Franca de Xira	77,5	Miranda do Douro	61,3
Almada	105,1	Oliveira do Hospital	76,8	Grândola	61,3
Manteigas	103,5	Freixo de Espada à Cinta	76,6	MINHO-LIMA	61,2
Oeiras	100,5	Crato	76,4	Alcanena	60,9
Caldas da Rainha	100,4	BAIXO VOUGA	76,2	Penela	60,7
<b>PORTUGAL CONTINENTAL</b>	100,0	Anadia	75,8	Estarreja	60,3
Portimão	98,2	ALTO TRÁS-OS-MONTES	75,1	Alvaiázere	59,8
Mirandela	97,9	BAIXO ALENTEJO	74,7	Vila Real de Santo António	59,8
BEIRA INTERIOR SUL	97,0	<b>ALGARVE</b>	74,6	Estremoz	59,5
Viana do Castelo	96,2	Santiago do Cacém	73,3	Oliveira de Azeméis	59,3
Penacova	95,6	Esposende	72,4	Miranda do Corvo	59,3
BEIRA INTERIOR NORTE	95,3	Cantanhede	72,3	Tábua	59,1
Chaves	94,9	Vila Nova de Famalicão	71,8	Serpa	59,0
Abrantes	94,2	Espinho	71,3	Arronches	58,5
Elvas	94,1	Castelo de Vide	70,2	Trancoso	58,4
Figueira da Foz	93,7	Viana do Alentejo	69,6	Tomar	58,2
COVA DA BEIRA	93,2	Vila Nova de Poiares	69,4	Fafe	57,9
Ansião	92,0	DOURO	69,3	Portel	57,6
<b>CENTRO</b>	91,4	PINHAL INTERIOR NORTE	69,3	Figueira de Castelo Rodrigo	57,4
Gouveia	89,9	Almeida	69,2	Pinhel	57,3
Cascais	89,0	Mora	69,0	Aguiar da Beira	57,1
Leiria	88,0	Mealhada	68,4	Castro Marim	57,0
ALTO ALENTEJO	88,0	MÉDIO TEJO	68,0	Vila do Bispo	57,0
Póvoa de Varzim	86,9	Penafiel	67,9	Penamacor	56,6
<b>NORTE</b>	86,8	Barcelos	67,9	Torre de Moncorvo	56,4
Macedo de Cavaleiros	85,1	Vila Nova de Gaia	67,3	Peniche	56,3

## Índice de Saúde e Assistência Social (continuação)

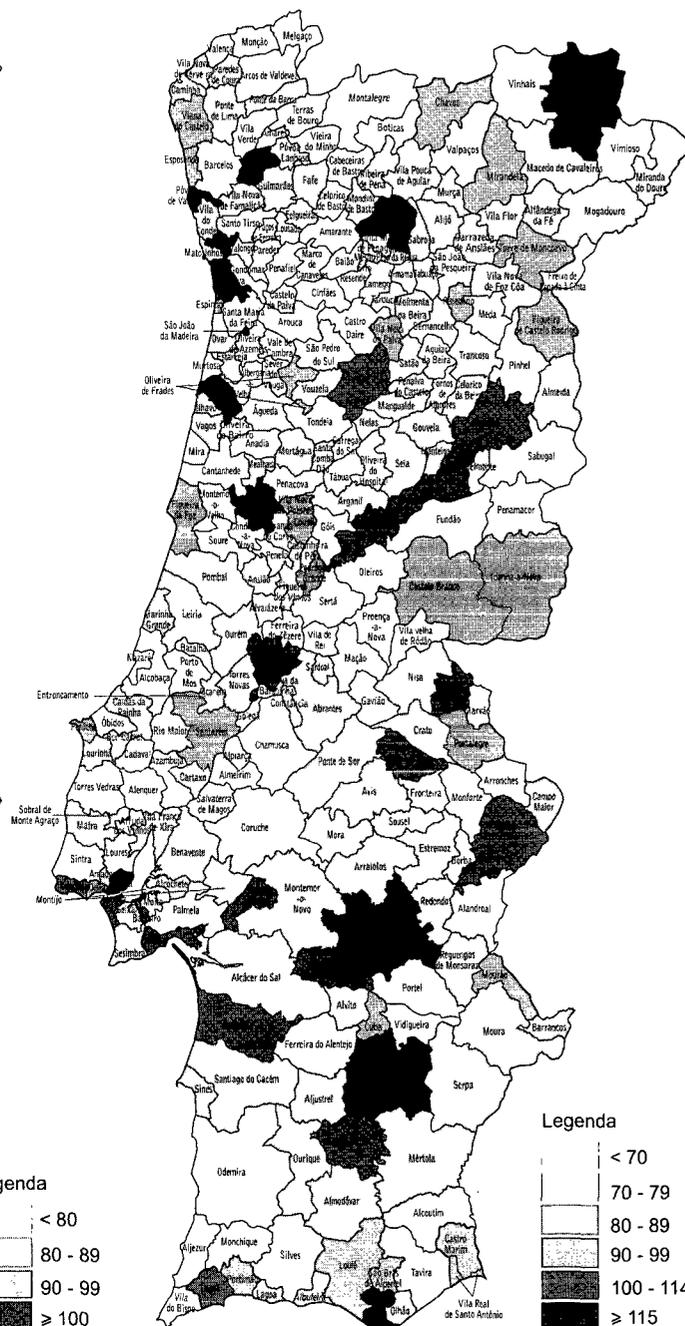
Ourique	56,2	Albufeira	46,2	Batalha	40,5
ENTRE DOURO E VOUGA	56,0	Almeirim	46,0	Sardoal	39,9
Nazaré	55,5	Mogadouro	46,0	Benavente	39,8
Vendas Novas	55,5	Proença-a-Nova	45,7	Mondim de Basto	39,8
Ferreira do Zêzere	55,2	Castelo de Paiva	45,6	Albergaria-a-Velha	39,7
Avis	55,1	Reguengos de Monsaraz	45,6	Oleiros	39,5
Tondela	54,9	Mafra	45,6	Sernancelhe	39,3
Marvão	54,8	Vila Nova da Barquinha	45,3	Gondomar	38,9
Vila do Conde	54,5	Tavira	45,1	Maia	38,7
Amarante	54,4	Loures	45,0	Alenquer	38,6
Carrizosa de Ansiães	54,3	Alandroal	45,0	Coruche	38,5
Alter do Chão	54,3	Lagoa	44,9	Cabeceiras de Basto	38,2
Lousã	53,9	Idanha-a-Nova	44,6	Póvoa de Lanhoso	38,0
Gavião	53,5	Vila Nova de Cerveira	44,5	Vinhais	38,0
Ílhavo	53,3	Moita	44,5	Baião	37,6
Golegã	53,1	Nelas	44,4	Vila Nova de Paiva	37,3
Alcochete	53,0	Ponte de Lima	44,4	Barrancos	37,1
Sousel	52,9	Cadaval	44,3	Vila Verde	36,7
Cuba	52,6	Vila de Rei	44,3	Bombarral	36,6
Arruda dos Vinhos	52,5	Chamusca	44,2	Tabuaço	36,4
Murtosa	52,4	Arcos de Valdevez	44,1	Arouca	36,3
Sesimbra	52,3	Mira	43,9	Mação	36,3
Pedrógão Grande	52,3	Mortágua	43,8	Rio Maior	35,0
Sobral de Monte Agraço	52,2	Figueiró dos Vinhos	43,8	Odemira	34,9
Santa Marta de Penaguião	52,2	Almodôvar	43,6	Oliveira de Frades	34,8
Terras de Bouro	51,9	Silves	43,6	Montalegre	34,7
Aljustrel	51,3	Paços de Ferreira	43,6	Armamar	34,5
Santa Maria da Feira	51,1	Amares	43,3	Valpaços	34,4
Ponte de Sôr	50,7	Penedono	43,0	Vale de Cambra	34,2
Santo Tirso	50,6	Moura	43,0	Valença	33,9
Mesão Frio	50,6	Monchique	42,9	Boticas	33,8
Caminha	50,5	Salvaterra de Magos	42,9	Monção	33,7
Meda	50,0	Aljô	42,7	Vouzela	33,4
Fornos de Algodres	49,9	Porto de Mós	42,5	Óbidos	33,2
Sintra	49,9	Sertã	42,5	Tarouca	33,0
Murça	49,9	Marco de Canavezes	42,5	Aljezur	32,9
Cartaxo	49,4	Borba	42,5	Moimenta da Beira	32,6
Vagos	49,1	TÂMEGA	42,4	São Pedro do Sul	32,4
Mourão	48,8	Mértola	42,2	Paredes	32,2
Sabrosa	48,8	Vila Pouca de Aguiar	41,9	Felgueiras	30,7
Campo Maior	48,1	Ferreira do Alentejo	41,7	Lousada	30,3
Valongo	48,1	PINHAL INTERIOR SUL	41,7	Sátão	30,3
Ourém	48,1	Redondo	41,5	São João da Pesqueira	30,1
Alpiarça	47,9	Santa Comba Dão	41,5	Penalva do Castelo	29,9
Pombal	47,7	Sever do Vouga	41,2	Carregal do Sal	29,3
Palmela	47,4	Mangualde	41,1	Cinfães	29,1
Marinha Grande	47,0	Seixal	41,1	Resende	28,2
Lourinhã	46,7	Vidigueira	40,9	Castro Daire	28,2
Ribeira de Pena	46,6	Celorico de Basto	40,9		
Castro Verde	46,6	Loulé	40,8		
Olhão	46,5	Melgaço	40,6		
Vieira do Minho	46,4	Ponte da Barca	40,5		

## ÍNDICE DE EDUCAÇÃO E CULTURA (Portugal Continental = 100)

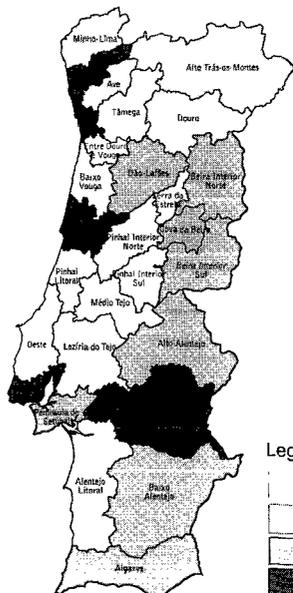
NUTS II



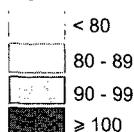
CONCELHOS



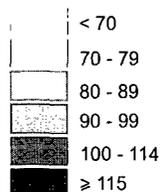
NUTS III



Legenda



Legenda



Índice de Educação e Cultura					
Lisboa	230,4	Vila Nova de Poiares	94,3	Ponte de Sôr	86,7
Coimbra	185,6	Vila Real de Santo António	93,7	Celorico da Beira	86,7
Porto	173,6	Oliveira de Frades	93,6	Gouveia	86,6
Évora	156,6	Santarém	93,5	Aljustrel	86,5
Braga	133,7	Portimão	93,3	Montemor-o-Novo	86,4
Vila Nova de Gaia	131,9	BEIRA INTERIOR SUL	93,2	MÉDIO TEJO	85,8
São João da Madeira	128,8	Idanha-a-Nova	93,2	Rio Maior	85,7
GRANDE LISBOA	127,2	Mirandela	93,2	Mértola	85,6
BAIXO MONDEGO	125,7	Figueira de Castelo Rodrigo	92,9	ALENTEJO LITORAL	85,4
Maia	123,8	Pedrogão Grande	92,9	Aljód	85,3
Beja	122,9	PENÍNSULA DE SETÚBAL	92,7	Vila Flor	85,3
Aveiro	122,3	ALGARVE	92,3	Seia	85,0
Faro	120,1	Vila Nova de Paiva	92,0	Mafra	84,9
Bragança	119,0	Penedono	91,8	Vila Nova de Cerveira	84,9
GRANDE PORTO	117,4	Chaves	91,6	Marinha Grande	84,9
Vila Real	116,6	Peniche	91,5	SERRA DA ESTRELA	84,8
Póvoa de Varzim	114,8	BEIRA INTERIOR NORTE	91,1	Seixal	84,7
Castro Verde	111,9	Mourão	91,1	PINHAL LITORAL	84,6
Viseu	111,6	COVA DA BEIRA	90,9	Olhão	84,5
Guarda	110,3	Oeiras	90,8	Lagoa	84,4
LISBOA E VALE DO TEJO	109,0	DÃO-LAFÕES	90,4	Pombal	84,4
Viana do Alentejo	108,6	Viana do Castelo	90,2	Paredes de Coura	84,3
ALENTEJO CENTRAL	108,6	Castro Marim	90,1	Tarouca	84,1
Cascais	108,0	ALTO ALENTEJO	89,9	Felgueiras	84,1
Castelo de Vide	107,1	Esposende	89,6	Murça	84,0
Grândola	106,3	Cuba	89,6	Montijo	83,8
Lagos	105,9	Sines	89,4	Miranda do Douro	83,8
Setúbal	104,9	Almodôvar	89,2	Sardoal	83,7
Almada	103,7	Sobral de Monte Agraço	89,0	Ansião	83,7
Vila Viçosa	103,4	Trancoso	88,9	Matosinhos	83,6
Alter do Chão	103,4	Crato	88,9	Barrancos	83,5
Entroncamento	102,6	ALTO TRÁS-OS-MONTES	88,8	Vila Franca de Xira	83,4
Barreiro	102,6	Arganil	88,7	Torres Novas	83,2
Elvas	102,2	DOURO	88,6	Bombarral	83,1
CÁVADO	100,6	Leiria	88,5	MINHO-LIMA	83,0
Vendas Novas	100,6	Arraiolos	88,4	PINHAL INTERIOR NORTE	82,9
PORTUGAL CONTINENTAL	100,0	Nelas	88,4	Miranda do Corvo	82,8
Tomar	99,8	Ribeira de Pena	88,3	Almeida	82,7
Covilhã	99,7	Moura	88,3	Águeda	82,7
Pampilhosa da Serra	99,5	Caldas da Rainha	88,3	Alvito	82,6
Castelo Branco	97,8	Caminha	88,1	Campo Maior	82,4
Lousã	97,3	Carrizada de Ansiães	88,1	Alfândega da Fé	82,4
ALENTEJO	96,8	Serpa	88,0	Ponte da Barca	82,3
Figueira da Foz	96,5	Albufeira	87,9	Alcanena	82,3
BAIXO ALENTEJO	96,1	Macedo de Cavaleiros	87,8	Constância	82,1
Torre de Moncorvo	95,9	Alcácer do Sal	87,4	Santiago do Cacém	82,1
Portalegre	95,8	BAIXO VOUGA	87,4	Figueiró dos Vinhos	81,9
Loulé	95,6	Estremoz	87,3	Mogadouro	81,9
São Brás de Alportel	95,5	Vila Nova de Foz Côa	87,1	Sabrosa	81,8
NORTE	95,4	Amadora	86,9	Estarreja	81,8
Espinho	95,1	Moimenta da Beira	86,9	Vouzela	81,6
CENTRO	94,8	Reguengos de Monsaraz	86,7	Santa Comba Dão	81,5

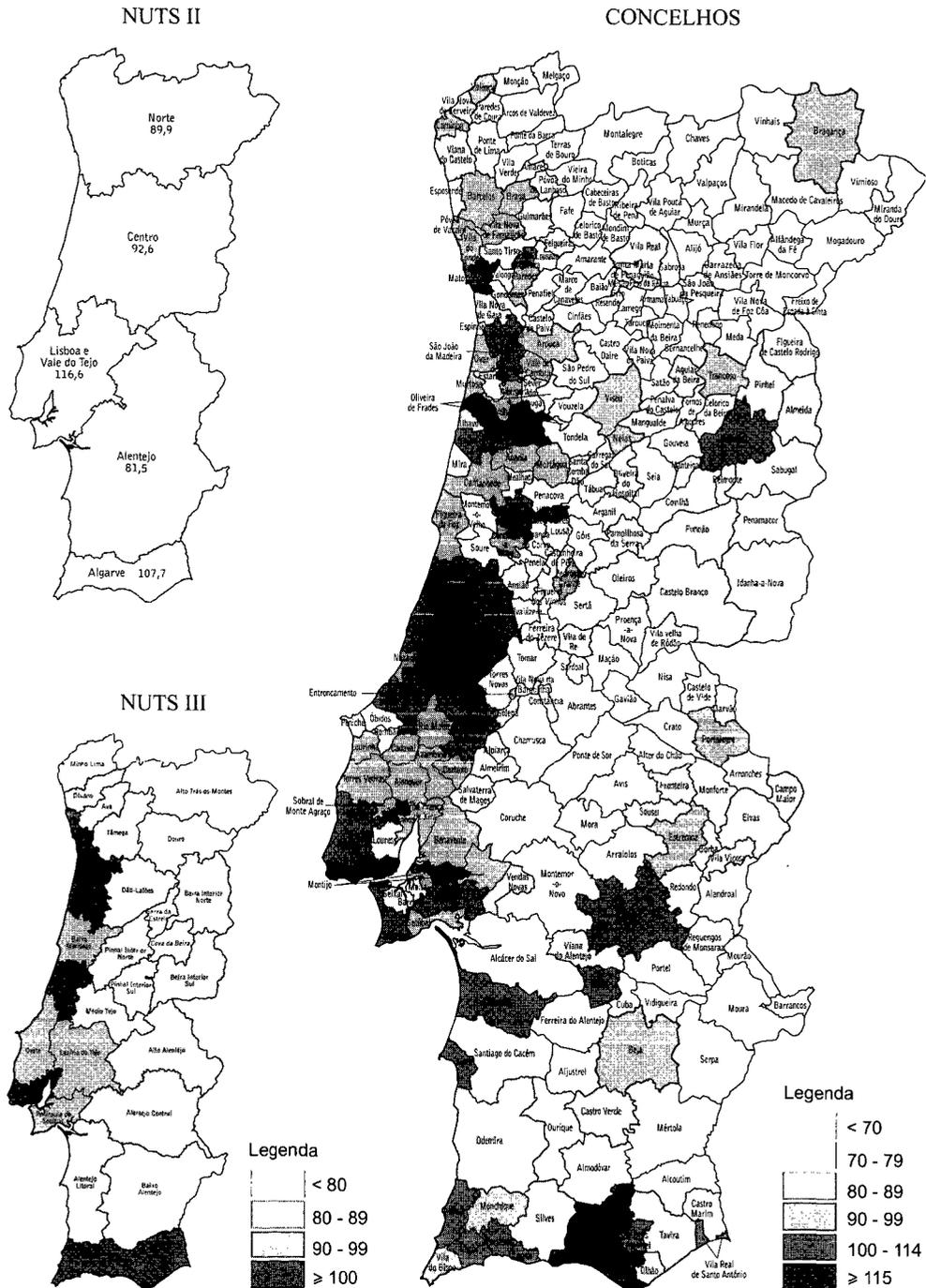
Índice de Educação e Cultura (continuação)					
Nisa	81,5	Celorico de Basto	78,2	Sousel	74,0
Lamego	81,5	Lousada	78,2	Penela	73,9
Alcochete	81,4	Condeixa-a-Nova	78,1	Monção	73,8
Peso da Régua	81,3	Vale de Cambra	78,1	Boticas	73,8
Ourém	81,3	Melgaço	78,0	Terras de Bouro	73,8
Mesão Frio	81,1	Cantanhede	77,9	Arcos de Valdevez	73,7
Ourique	81,1	Santo Tirso	77,9	Azambuja	73,6
Sesimbra	81,1	Sertã	77,8	Cadaval	73,4
Ovar	81,1	Torres Vedras	77,7	Almeirim	73,4
LEZÍRIA DO TEJO	81,0	Tábua	77,7	Batalha	73,4
Vila do Conde	80,9	Chamusca	77,7	Meda	73,2
Vagos	80,9	Sintra	77,7	Marco de Canavezes	73,1
Manteigas	80,7	Valença	77,6	Anadia	73,0
Fafe	80,6	ENTRE DOURO E VOUGA	77,6	Montemor-o-Velho	72,9
Mortágua	80,3	Montalegre	77,5	Vila Velha de Ródão	72,6
Sátão	80,3	Albergaria-a-Velha	77,4	Oliveira do Bairro	72,5
Póvoa de Lanhoso	80,3	Proença-a-Nova	77,4	Portel	72,5
Cartaxo	80,3	Vidigueira	77,3	Monchique	72,5
Vila Nova de Famalicão	80,2	Vila de Rei	77,2	Mação	72,4
Guimarães	80,2	Paredes	77,0	Sabugal	72,3
Avis	80,1	Cinfães	77,0	Tabuaço	72,0
Ílhavo	80,0	Resende	77,0	Ferreira do Zêzere	71,9
Abrantes	80,0	TÁMEGA	76,9	Salvaterra de Magos	71,9
Benavente	79,9	Armamar	76,8	Amarante	71,8
Lourinhã	79,9	São Pedro do Sul	76,5	Alvaiázere	71,8
Mondim de Basto	79,7	Monforte	76,4	Alpiarça	71,2
Ponte de Lima	79,7	Góis	76,4	Sernancelhe	71,2
AVE	79,7	Barcelos	76,4	Coruche	71,1
Vila Pouca de Aguiar	79,7	Loures	76,3	Castanheira de Pera	71,0
Penalva do Castelo	79,6	São João da Pesqueira	76,3	Santa Maria da Feira	70,9
Mangualde	79,6	Pinhel	76,2	Penamacor	70,6
Belmonte	79,6	Ferreira do Alentejo	76,2	Vimioso	70,6
Mealhada	79,6	Tavira	76,1	Marvão	70,3
Redondo	79,5	Vila Nova da Barquinha	76,1	Gondomar	70,1
OESTE	79,4	Mira	76,0	Baião	70,0
Golegã	79,4	Silves	75,9	Soure	69,8
Tondela	79,3	Odemira	75,7	Mora	69,7
Valongo	79,2	Oliveira do Hospital	75,5	Oleiros	69,2
Palmela	79,1	PINHAL INTERIOR SUL	75,4	Freixo de Espada à Cinta	69,2
Arouca	79,0	Aguiar da Beira	75,3	Alcoutim	69,1
Fundão	79,0	Oliveira de Azeméis	75,3	Aljezur	68,3
Carregal do Sal	79,0	Alenquer	75,1	Nazaré	67,5
Castelo de Paiva	78,9	Vinhais	74,9	Óbidos	66,4
Paços de Ferreira	78,8	Penafiel	74,9	Santa Marta de Penaguião	65,9
Fornos de Algodres	78,7	Vila Verde	74,8	Alandroal	65,2
Sever do Vouga	78,7	Moita	74,8	Arruda dos Vinhos	64,4
Castro Daire	78,6	Alcobaça	74,8	Borba	63,5
Fronteira	78,6	Valpaços	74,7	Vila do Bispo	62,0
Vieira do Minho	78,6	Penacova	74,6		
Cabeceiras de Basto	78,4	Gavião	74,4		
Amares	78,4	Murtosa	74,3		
Arronches	78,3	Porto de Mós	74,2		



Índice de Rendimento					
Lisboa	184,8	São Brás de Alportel	95,8	Alcanena	86,9
Oeiras	142,6	Ílhavo	95,6	Santiago do Cacém	86,6
Porto	140,2	Mafra	95,6	ALTO ALENTEJO	86,2
GRANDE LISBOA	138,4	BAIXO VOUGA	94,7	Miranda do Douro	86,1
Cascais	134,8	Tomar	94,6	Borba	86,1
São João da Madeira	125,8	Bragança	93,9	BEIRA INTERIOR SUL	86,0
Entroncamento	123,9	Benavente	93,7	ALENTEJO LITORAL	86,0
Almada	121,5	Mealhada	93,7	Elvas	86,0
<b>LISBOA E VALE DO TEJO</b>	120,5	Gondomar	93,2	Ourém	85,9
Coimbra	119,7	Vendas Novas	93,0	<b>ALENTEJO</b>	85,6
Sintra	119,4	Águeda	92,9	Olhão	85,4
Faro	119,3	Albergaria-a-Velha	92,6	Vila do Conde	85,1
Aveiro	116,9	Castelo Branco	92,6	Caminha	84,6
Barreiro	116,1	Vila Real de Santo António	92,4	Santo Tirso	84,5
Setúbal	114,8	Lagoa	92,2	Santa Maria da Feira	84,4
Amadora	113,9	Torres Novas	92,2	Cantanhede	84,3
Sines	113,2	Alenquer	92,0	Soure	84,2
Marinha Grande	111,6	Condeixa-a-Nova	91,8	Covilhã	83,8
Matosinhos	110,8	Azambuja	91,8	Póvoa de Varzim	83,8
Vila Franca de Xira	110,3	Vila Real	91,1	Monchique	83,4
<b>PENÍNSULA DE SETÚBAL</b>	110,3	Viana do Castelo	91,0	Reguengos de Monsaraz	83,3
<b>GRANDE PORTO</b>	109,0	Torres Vedras	90,9	Avis	83,3
Seixal	108,2	Sobral de Monte Agraço	90,8	Castelo de Vide	83,1
Loures	107,0	Moita	90,7	Mangualde	82,9
Maia	106,4	<b>MÉDIO TEJO</b>	90,7	Mogadouro	82,6
Portimão	105,9	Constância	90,5	Guimarães	82,5
Palmela	104,8	Estremoz	90,3	<b>CÁVADO</b>	82,4
Vila Nova de Gaia	103,9	Cartaxo	90,2	Silves	82,4
Leiria	103,4	Vila Viçosa	90,1	<b>AVE</b>	82,4
Évora	103,4	Arruda dos Vinhos	90,1	Nelas	82,3
Alcochete	103,3	Rio Maior	90,0	Fronteira	82,2
Albufeira	102,5	Montemor-o-Novo	90,0	Torre de Moncorvo	82,2
Portalegre	101,1	Lousã	89,5	Coruche	82,0
Montijo	100,9	Castro Verde	89,5	Chaves	81,9
<b>BAIXO MONDEGO</b>	100,9	<b>CENTRO</b>	89,2	Miranda do Corvo	81,9
<b>PORTUGAL CONTINENTAL</b>	100,0	Anadia	88,9	Pombal	81,7
Beja	99,9	<b>ALENTEJO CENTRAL</b>	88,8	Nazaré	81,5
Figueira da Foz	99,9	Vila Velha de Ródão	88,7	Porto de Mós	81,4
Loulé	99,1	Oliveira de Azeméis	88,6	Mourão	81,4
Braga	98,9	<b>NORTE</b>	88,6	Bombarral	81,1
Sesimbra	98,8	Batalha	88,5	<b>BEIRA INTERIOR NORTE</b>	81,0
Lagos	98,3	<b>OESTE</b>	88,4	Ansião	81,0
Caldas da Rainha	98,0	Vale de Cambra	88,1	Ponte de Sôr	81,0
Espinho	97,8	Vila Nova de Famalicão	87,9	<b>COVA DA BEIRA</b>	80,7
Estarreja	97,5	Ovar	87,7	<b>DÃO-LAFÕES</b>	80,5
Viseu	96,9	Abrantes	87,7	<b>BAIXO ALENTEJO</b>	80,4
Guarda	96,7	<b>LEZÍRIA DO TEJO</b>	87,6	Almeirim	80,0
Valongo	96,7	<b>ENTRE DOURO E VOUGA</b>	87,4	Peso da Régua	79,9
<b>PINHAL LITORAL</b>	96,6	Vila Nova da Barquinha	87,4	Alcácer do Sal	79,9
Campo Maior	96,5	Alcobaça	87,2	Sardoal	79,8
Santarém	96,2	Oliveira do Bairro	87,1	Aljustrel	79,8
<b>ALGARVE</b>	95,8	Grândola	87,0	Tavira	79,7

<b>Índice de Rendimento (continuação)</b>					
Peniche	79,6	Arronches	74,4	Marco de Canavezes	69,1
Lourinhã	79,2	Celorico da Beira	74,3	Góis	69,0
Nisa	78,9	Pedrógão Grande	74,2	Tábua	68,9
Marvão	78,8	Ferreira do Alentejo	74,0	Redondo	68,8
Golegã	78,8	Esposende	73,9	Vimioso	68,8
Vila Nova de Poiares	78,6	Vila Nova de Cerveira	73,6	Figueira de Castelo Rodrigo	68,8
MINHO-LIMA	77,8	Odemira	73,6	Ourique	68,3
Castanheira de Pera	77,7	Crato	73,5	Mértola	68,3
Belmonte	77,7	Ponte da Barca	73,4	Alijó	68,3
Montemor-o-Velho	77,6	Mortágua	73,4	Amares	68,0
Mirandela	77,4	Valença	73,4	Penamacor	68,0
Lamego	77,4	Mação	73,1	São João da Pesqueira	67,3
Mora	77,2	Fafe	73,0	Lousada	67,3
Figueiró dos Vinhos	77,0	Vieira do Minho	72,9	Ponte de Lima	67,2
Viana do Alentejo	76,9	Manteigas	72,8	Sátão	67,2
Salvaterra de Magos	76,8	Sertã	72,7	Santa Marta de Penaguião	66,8
Monforte	76,8	Fornos de Algodres	72,2	Carraceda de Ansiães	66,7
Seia	76,8	Alvito	72,0	Alfândega da Fé	66,7
PINHAL INTERIOR NORTE	76,6	Cuba	71,9	Aguiar da Beira	66,7
Felgueiras	76,6	Meda	71,7	Castro Daire	66,4
Penedono	76,6	São Pedro do Sul	71,7	Vila Pouca de Aguiar	66,1
Trancoso	76,4	Pinhel	71,7	Arcos de Valdevez	66,1
Alter do Chão	76,4	Oliveira do Hospital	71,5	Murça	65,8
ALTO TRÁS-OS-MONTES	76,3	Mesão Frio	71,5	Barrancos	65,7
Alvaiázere	76,2	Vila de Rei	71,3	Idanha-a-Nova	65,5
Gouveia	76,2	Paços de Ferreira	71,1	Moimenta da Beira	64,7
Fundão	76,2	Arraiolos	70,9	Paredes de Coura	64,6
Murtosa	76,1	Gavião	70,8	Vila Flor	64,6
Montalegre	76,1	Penacova	70,8	Sabrosa	63,9
Penafiel	76,1	Monção	70,6	Alcoutim	63,6
Santa Comba Dão	76,1	Freixo de Espada à Cinta	70,6	Tabuaço	63,5
Sever do Vouga	76,1	Carregal do Sal	70,6	Portel	63,4
SERRA DA ESTRELA	76,1	PINHAL INTERIOR SUL	70,5	Terras de Bouro	63,4
Almeida	76,1	Almodôvar	70,4	Alandroal	63,3
Melgaço	76,0	Cabeceiras de Basto	70,3	Armamar	63,3
Vidigueira	76,0	Sabugal	70,2	Vila Verde	63,1
Moura	76,0	Vila Nova de Paiva	70,1	Sernancelhe	62,8
Arganil	75,8	Penela	70,1	Boticas	62,6
Vagos	75,7	Óbidos	69,9	Vinhais	62,6
Sousel	75,6	Vouzela	69,9	Oleiros	62,5
Cadaval	75,6	Amarante	69,9	Tarouca	62,2
Mira	75,5	Póvoa de Lanhoso	69,8	Resende	61,5
Vila Nova de Foz Côa	75,4	Oliveira de Frades	69,8	Baião	61,4
Pampilhosa da Serra	75,3	Castro Marim	69,8	Cinfães	61,0
Aljezur	75,2	TÂMEGA	69,7	Valpaços	60,6
Penalva do Castelo	74,8	Ferreira do Zêzere	69,6	Ribeira de Pena	60,3
DOURO	74,7	Chamusca	69,4	Celorico de Basto	59,5
Castelo de Paiva	74,7	Vila do Bispo	69,4	Mondim de Basto	58,1
Barcelos	74,6	Paredes	69,4		
Alpiarça	74,6	Proença-a-Nova	69,3		
Tondela	74,4	Arouca	69,2		
Macedo de Cavaleiros	74,4	Serpa	69,2		

## ÍNDICE DE EMPREGO E ACTIVIDADE ECONÓMICA (Portugal Continental = 100)



Índice de Emprego e Actividade Económica					
Lisboa	208,2	Guarda	101,1	Vila Nova de Gaia	89,1
Porto	153,8	<b>PORTUGAL CONTINENTAL</b>	100,0	Elvas	88,9
Albufeira	145,5	Caldas da Rainha	100,0	<b>MÉDIO TEJO</b>	88,9
São João da Madeira	141,9	Grândola	99,8	Seixal	88,6
<b>GRANDE LISBOA</b>	135,4	Torres Vedras	99,4	Vila Viçosa	88,6
Oeiras	134,5	Sobral de Monte Agraço	98,7	Oliveira do Hospital	88,6
Loulé	132,7	Nelas	98,5	<b>CÁVADO</b>	88,3
Batalha	125,3	Beja	98,5	Loures	88,2
Aveiro	124,8	Setúbal	98,2	Espinho	87,8
Faro	123,1	Azambuja	97,6	Mealhada	87,7
Leiria	120,7	<b>OESTE</b>	97,6	Ílhavo	87,4
Arruda dos Vinhos	119,0	Rio Maior	97,5	<b>ALENTEJO CENTRAL</b>	87,2
Maia	116,8	Anadia	97,3	Guimarães	87,2
<b>LISBOA E VALE DO TEJO</b>	116,6	Braga	97,0	Estarreja	87,0
Cascais	115,9	<b>BAIXO MONDEGO</b>	96,9	Vendas Novas	86,7
Águeda	115,7	Benavente	96,7	Tomar	86,6
Alcanena	115,5	Cantanhede	96,4	Castelo Branco	86,5
<b>PINHAL LITORAL</b>	114,8	Viseu	96,1	<b>BEIRA INTERIOR NORTE</b>	86,2
Sines	113,7	Nazaré	96,0	Almeirim	85,8
Ourém	113,5	Arouca	96,0	Campo Maior	85,8
Pombal	112,5	Montijo	95,4	Alpiarça	85,7
Coimbra	111,7	Trancoso	95,2	Silves	85,3
Palmela	111,3	Albergaria-a-Velha	94,8	Aguiar da Beira	85,3
Portimão	110,6	Lourinhã	94,6	Felgueiras	85,3
Vagos	109,8	Vale de Cambra	94,4	Barreiro	85,1
Oliveira de Azeméis	109,3	Mortágua	94,3	Torres Novas	84,9
Lagos	109,0	Vila Nova de Famalicão	93,8	Fronteira	84,6
Amadora	108,4	Caminha	93,8	<b>ALENTEJO LITORAL</b>	84,6
Vila Real de Santo António	108,3	Cartaxo	93,8	Borba	84,5
Alvito	108,3	<b>PENÍNSULA DE SETÚBAL</b>	93,7	Lousã	84,5
<b>ENTRE DOURO E VOUGA</b>	108,3	Cadaval	93,7	Ansião	84,5
Santa Maria da Feira	107,8	Paredes	93,3	Oliveira de Frades	84,4
<b>ALGARVE</b>	107,7	Vila Franca de Xira	93,0	<b>AVE</b>	84,4
Mafra	107,2	<b>CENTRO</b>	92,6	Santo Tirso	84,4
Sintra	106,6	Alenquer	92,6	Olhão	84,3
Vila Nova de Poiares	106,4	Vila do Conde	92,4	Vila Real	84,2
Lagoa	106,1	Barcelos	92,0	Almeida	83,9
<b>GRANDE PORTO</b>	105,7	Bragança	91,8	Viana do Castelo	83,8
Oliveira do Bairro	105,6	Entroncamento	91,8	Miranda do Corvo	83,7
Alcochete	104,9	Murtosa	91,7	Golegã	83,6
Évora	104,8	<b>LEZÍRIA DO TEJO</b>	91,7	Viana do Alentejo	83,5
Marinha Grande	104,2	Figueira da Foz	91,5	Bombarral	83,1
Alcobaça	103,6	Póvoa de Varzim	91,4	Lousada	82,1
Matosinhos	103,0	Ovar	91,1	Tavira	81,9
São Brás de Alportel	102,6	Portalegre	91,0	Belmonte	81,7
Porto de Mós	102,2	Condeixa-a-Nova	90,7	Vila de Rei	81,6
Almada	102,2	Vila do Bispo	90,2	<b>PINHAL INTERIOR NORTE</b>	81,5
Santarém	102,0	Estremoz	90,1	<b>ALENTEJO</b>	81,5
<b>BAIXO VOUGA</b>	101,9	Pedrógão Grande	90,0	Pinhel	81,4
Aljezur	101,9	<b>NORTE</b>	89,9	Alvaiázere	81,3
Sesimbra	101,2	Monchique	89,7	Sever do Vouga	81,2
Paços de Ferreira	101,2	Valença	89,6	Odemira	81,1

<b>Índice de Emprego e Actividade Económica (continuação)</b>					
BEIRA INTERIOR SUL	80,6	Amarante	74,0	Mourão	63,6
ALTO ALENTEJO	80,5	Reguengos de Monsaraz	73,4	Meda	63,5
Santa Comba Dão	80,1	Arraiolos	73,2	Penadono	63,2
DÃO-LAFÕES	80,0	BAIXO ALENTEJO	72,9	Serpa	63,1
Fundão	80,0	Góis	72,8	Alfândega da Fé	62,5
Castro Verde	79,9	Montemor-o-Velho	72,7	Armamar	62,1
Sousel	79,9	Santiago do Cacém	72,7	Castelo de Paiva	62,0
Pampilhosa da Serra	79,5	Chamusca	72,5	São Pedro do Sul	61,5
Sertã	79,4	Amares	72,3	Castro Marim	61,4
Gondomar	79,3	Alter do Chão	72,2	Barrancos	61,1
Óbidos	78,8	Peso da Régua	72,0	Moura	60,8
COVA DA BEIRA	78,6	São João da Pesqueira	71,8	Mértola	60,8
Salvaterra de Magos	78,4	Tondela	71,7	Vieira do Minho	60,7
Montemor-o-Novo	78,4	Arganil	71,6	Idanha-a-Nova	60,7
Ponte de Lima	78,4	Gouveia	71,3	Ferreira do Alentejo	60,4
MINHO-LIMA	78,3	Constância	71,3	Cuba	60,1
Mira	78,3	Melgaço	71,2	Paredes de Coura	60,0
Peniche	78,2	ALTO TRÁS-OS-MONTES	70,4	Aljustrel	59,9
Esposende	77,8	Moita	70,4	Terras de Bouro	59,5
Figueiró dos Vinhos	77,7	Ourique	70,2	Vila Pouca de Aguiar	59,3
Penacova	77,7	Vila Velha de Ródão	70,2	Moimenta da Beira	59,2
Penela	77,6	Tábua	70,1	Portel	59,2
Soure	77,3	Castro Daire	70,0	Torre de Moncorvo	58,8
Covilhã	77,2	Ferreira do Zêzere	69,5	Valpaços	58,1
Castelo de Vide	77,2	Fornos de Algodres	69,1	Mesão Frio	57,2
Abrantes	77,0	Celorico da Beira	69,1	Crato	57,1
Proença-a-Nova	77,0	SERRA DA ESTRELA	68,7	Sátão	56,5
Coruche	76,8	Vila Verde	68,6	Cinfães	55,8
Mora	76,7	Fafe	68,3	Tarouca	55,3
TÂMEGA	76,6	Seia	67,2	Murça	55,3
Monforte	76,6	Vila Nova de Foz Côa	67,2	Cabeceiras de Basto	54,9
Marco de Canavezes	76,6	Marvão	67,2	Sabrosa	53,7
Oleiros	76,5	Sardoal	67,0	Vila Flor	53,7
Ponte de Sôr	76,4	Póvoa de Lanhoso	67,0	Penalva do Castelo	53,5
Miranda do Douro	76,4	Vouzela	66,7	Sernancelhe	53,2
Alcácer do Sal	76,3	Alandroal	66,4	Vinhais	52,7
Redondo	76,2	Vimioso	66,4	Mondim de Basto	52,4
Figueira de Castelo Rodrigo	76,1	Mação	66,4	Vila Nova de Paiva	51,9
PINHAL INTERIOR SUL	76,0	Macedo de Cavaleiros	66,3	Boticas	51,6
Chaves	75,3	Ponte da Barca	66,0	Gavião	51,0
Mirandela	75,3	Vila Nova da Barquinha	65,9	Alijó	50,1
Valongo	75,1	Resende	65,9	Celorico de Basto	49,3
Montalegre	75,1	Vidigueira	65,8	Tabuaço	48,3
Penafiel	74,7	Castanheira de Pera	65,6	Ribeira de Pena	47,6
Penamacor	74,6	Nisa	65,5	Carrazeda de Ansiães	47,5
Carregal do Sal	74,6	Arronches	65,3	Baião	45,1
Mangualde	74,5	Alcúcutim	65,3	Santa Marta de Penaguião	44,4
Sabugal	74,5	Almodôvar	64,5	Freixo de Espada à Cinta	43,7
Monção	74,4	Lamego	64,3		
Vila Nova de Cerveira	74,3	Arcos de Valdevez	63,9		
Avis	74,2	DOURO	63,9		
Manteigas	74,2	Mogadouro	63,7		

## Utilização de bases de microdados na investigação em ciências sociais

Autores:

Rosário Mauritti

Susana da Cruz Martins

Ana Simões Antunes

António Firmino da Costa

**VOLUME II**

**2° QUADRIMESTRE DE 2002**

---

---

## UTILIZAÇÃO DE BASES DE MICRODADOS NA INVESTIGAÇÃO EM CIÊNCIAS SOCIAIS

---

---

---

---

## THE UTILIZATION OF MICRO-DATABASES IN SOCIAL SCIENCES RESEARCH

---

---

Autores: Rosário Mauritti

- Investigadora do Centro de Investigação e Estudos de Sociologia do Instituto Superior de Ciências do Trabalho e da Empresa (CIES/ISCTE).

Susana da Cruz Martins

- Investigadora do CIES/ISCTE.

Ana Simões Antunes

- Departamento de Estatísticas Sociais, Coordenadora do Núcleo de Estatísticas do Emprego do Serviço de Estatísticas do Trabalho, Instituto Nacional de Estatística.

António Firmino da Costa

- Professor Auxiliar do ISCTE. Presidente da Direcção e Coordenador Científico do CIES/ISCTE.

### RESUMO:

- Este artigo procura analisar as possibilidades e limites da utilização de microdados estatísticos produzidos pelo INE na investigação em ciências sociais. Evocam-se a este propósito questões centrais que se colocam no uso desta informação, tendo em consideração, por um lado, o princípio fundamental de salvaguarda do segredo estatístico e, por outro, as potencialidades inscritas no acesso às bases de microdados. Reflecte-se assim sobre os procedimentos de anonimização, as condições de disponibilização da metainformação, a validação e o controlo de qualidade dos dados, designadamente no que respeita à harmonização e à comparabilidade, e sobre a concertação de critérios de divulgação da informação.

### PALAVRAS-CHAVE:

- *Microdados, anonimização, metainformação, harmonização, comparabilidade e divulgação dos dados.*

### ABSTRACT:

- The present article analyses the possibilities and limitations of the utilization of statistical microdata (produced by the Portuguese National Statistics Institute) in social sciences research. In order to do this, we aim to consider the fundamental principle of statistic privacy, on one hand, and the possibilities to research offered by the access to micro-databases, on the other. Thus, we reflect on the anonimiza-

tion procedures, on the conditions of reception of metadata, on the validation and control of data, namely harmonization and comparability, and on the criteria of information's diffusion.

**KEY-WORDS:**

- *Micro-data, anonimization, meta-information, harmonization, comparability and diffusion of data.*

---

## 1. ACESSO A BASES DE MICRODADOS ESTATÍSTICOS: POTENCIALIDADES E PROBLEMAS

---

A utilização de dados estatísticos produzidos institucionalmente é uma das peças-chave da investigação em ciências sociais. Foi assim desde os primeiros passos do desenvolvimento sistemático destas ciências, ainda no século XIX.

Mas, hoje, o recurso a indicadores estatísticos tornou-se ainda muito mais importante, nomeadamente em áreas científicas como a sociologia, a economia, a demografia e a geografia. Isso acontece, aliás, tanto na investigação fundamental, onde tem permitido análises decisivas para o avanço das ciências sociais, como na investigação aplicada, nomeadamente na que tem visado apoiar processos de planeamento, de decisão e de avaliação, em diversos domínios das políticas públicas.

É bem conhecido nas ciências sociais que nem só a informação estatística – nem sequer, de um modo mais geral, só a informação empírica quantitativa – é susceptível de apoiar a produção de conhecimento substantivo sobre as relações sociais, as suas configurações e as suas dinâmicas. Também a informação qualitativa decorrente de observação controlada, nas variadas formas em que as ciências sociais a utilizam, proporciona base empírica relevante para esse conhecimento.

Reconhecê-lo não equivale, porém, a qualquer subalternização dos dados quantitativos na caracterização da realidade social, nem desse género específico de dados quantitativos que são os indicadores estatísticos institucionais.

Por vezes, alguns investigadores têm contraposto, em regime de exclusão recíproca, a pertinência da informação quantitativa ou da informação qualitativa para a produção de enunciados cognitivos válidos e significativos sobre a realidade social.

No entanto, pelo menos na sociologia, as posições mais consistentes analiticamente, e mais cognitivamente produtivas, têm adoptado uma perspectiva de pluralismo metodológico, para a qual as diferenças entre informação empírica de carácter quantitativo e qualitativo não são de essência mas de procedimento. O diálogo reflectido entre métodos e a integração judiciosa de fontes diversas de informação empírica tem-se revelado crucial para o avanço das ciências sociais.

Como referido, a informação estatística institucionalmente produzida, se sempre foi importante para a investigação em ciências sociais, adquiriu hoje uma presença muito maior e mais diversificada. E, do mesmo passo, adquiriu também níveis de exigência muito mais elevados. Isso aconteceu em correspondência, por um lado, com o desenvolvimento interno destas ciências, e, por outro, com os desenvolvimentos da própria sociedade.

Deste ponto de vista, multiplicou-se exponencialmente o pedido social de informação estatística, sobre os mais diversos aspectos. Tipicamente, essa procura visa a produção de caracterizações e a análise de distribuições, relações e evoluções, respeitantes a indivíduos e grupos, a entidades e agregados, a atributos e actividades, a comportamentos e opiniões.

Mas para além disso, a informação estatística passou, ela própria, a fazer parte constitutiva importante de vários processos sociais. É um dos aspectos em que se manifesta a reflexividade cognitiva e institucional das sociedades contemporâneas, hoje em dia objecto de análise incontornável das ciências sociais, e tematizada publicamente sob expressões como “sociedade da informação” ou “sociedade do conhecimento”, associada a processos como os de “globalização” ou de “inovação tecnológica”, de “regulação” ou de “governança”.

Com efeito, a presença social da informação estatística institucional verifica-se, hoje, de forma intensiva, no plano de sistemas como o político-administrativo, o económico-empresarial, o educacional-formativo ou o comunicacional-mediático. Mas repercute-se, também, no plano dos modos de vida quotidiana, no dos processos de acção colectiva e no dos direitos, liberdades e garantias.

A informação estatística surge, pois, na intersecção de duas vertentes decisivas da constituição das sociedades contemporâneas:

- mecanismos importantíssimos de reflexividade cognitiva e instrumentalidade social;

mas, igualmente,

- preocupações cívicas com os direitos e as liberdades, nomeadamente no que concerne à defesa da privacidade e da autonomia individuais.

Do ponto de vista dos desenvolvimentos das ciências sociais, a evolução do seu objecto de estudo, isto é, da sociedade, com as tendências apontadas, bem como a crescente informação estatística institucionalmente produzida e, ainda, a disponibilidade de novas tecnologias de armazenagem e tratamento de informação, incomparavelmente mais poderosas do que há alguns anos atrás, confluem com a recolocação dos objectivos de investigação e dos procedimentos de análise em patamares muito mais exigentes.

As análises convencionais de dados agregados, habitualmente fornecidas pelas instituições oficiais de produção de informação estatística, sem que tenham perdido utilidade, tornaram-se claramente insuficientes.

Muitas das interrogações analíticas pertinentes nas ciências sociais actuais, e certamente alguns dos problemas de investigação mais interessantes e que mais podem fazer avançar o conhecimento científico sobre as configurações de sociedade e as dinâmicas sociais contemporâneas, requerem hoje, para que se possa investigar com vista à respectiva elucidação, a utilização de bases de microdados estatísticos institucionalmente produzidos.

O trabalho de investigação conduzido com recurso directo aos microdados possibilita:

- *análises de dados mais finas;*
- a construção de *variáveis compostas de tipos diversificados*, mais directamente ajustadas a parâmetros teóricos e temas de investigação;

- a realização de *análises multivariadas*, estabelecendo relações entre um grande número de indicadores;
- um modo de *relação experimental com as bases de dados*, permitindo um número alargado de tentativas sucessivas de análise, conduzidas segundo um processo de interacção entre parâmetros teóricos, ensaios de operacionalização e tentativas de interpretação de resultados.

Em simultâneo, as ciências sociais não podem deixar de se colocar na primeira linha das preocupações cívicas e éticas com a defesa dos direitos, as liberdades e as garantias dos cidadãos, tal como se configuram numa sociedade democrática contemporânea.

O conhecimento aprofundado que têm dos processos sociais atribui-lhes responsabilidades acrescidas, mais ainda quando está em causa algo que se situa no seu próprio âmbito, isto é, os dispositivos de recolha de informação sistemática e produção de conhecimento analítico sobre a realidade social. Os investigadores destas áreas, não só têm obrigação de respeitar as disposições legalmente estabelecidas a este propósito, como estão provavelmente melhor colocados do que muitos outros para detectar riscos e sugerir soluções aperfeiçoadas a este respeito.

É também neste sentido que ganha particular importância o carácter institucional da informação estatística aqui em causa. O Sistema Estatístico Nacional e, em particular, o Instituto Nacional de Estatística (INE) têm aqui uma responsabilidade particular quanto à possibilidade de supervisão técnica e responsabilização pública, sob tutela com legitimidade democrática, destes processos de recolha e disponibilização de informação estatística fina, sistemática e abrangente.

---

## 2. FONTES DE INFORMAÇÃO ESTATÍSTICA: NOVAS POSSIBILIDADES DO TRATAMENTO DE MICRODADOS

---

Está em desenvolvimento uma nova fase de interligação entre os produtores institucionais de informação estatística (em especial, o INE) e a comunidade científica, nomeadamente as equipas de investigadores em ciências sociais.

Neste contexto, procura-se a implementação de um novo quadro institucional que possibilite a divulgação regulamentada de microdados estatísticos. O protocolo entre o Ministério da Ciência e da Tecnologia (MCT) e o Instituto Nacional de Estatística, em 1999, estabelece esse novo quadro.

Antes de passar à análise detalhada do protocolo, importa ainda explicitar alguns dos vectores que poderão contribuir para o desenvolvimento de uma prática testada, consolidada e produtiva desses novos moldes de interligação. Uma forma de o

concretizar será certamente através da implementação de projectos de investigação científica.<sup>11</sup> Estes poderão contribuir para isso em termos de:

- *procedimentos técnicos* (suportes tecnológicos e processamento informático das bases de dados);
- *metainformação estatística* (sistematização, actualização e aperfeiçoamento permanentes, susceptíveis de serem conseguidos através da respectiva utilização na investigação);
- *processos de anonimização* (soluções exigentes quanto a garantias de anonimato e confidencialidade e, simultaneamente, viabilizadoras de aprofundamentos analíticos);
- *controlo dos dados e resultados a divulgar* (quanto a aspectos como validade, representatividade, fiabilidade e significância);
- *comparabilidade* (entre indicadores nacionais e internacionais e entre fontes estatísticas);
- *aperfeiçoamento* das operações estatísticas e aproveitamento analítico das suas potencialidades (testes e sugestões a partir das experiências de investigação);
- *processos de trabalho* (entre técnicos do INE e equipas de investigadores em ciências sociais).

A possibilidade de tratamento de microdados recolhidos e processados para fins estatísticos por uma instituição pública como o INE, constitui uma novidade de grande importância para o desenvolvimento das actividades de investigação em ciências sociais. Quais as razões que fundamentam a disponibilização destes dados para fins de investigação científica?

No quadro convencional tradicional, quando pretendem desenvolver uma investigação com apoio de estatísticas produzidas institucionalmente, os investigadores têm ao seu dispor:

- publicações e relatórios disponíveis ao público em geral;
- informação divulgada em linha na Internet;
- a possibilidade de solicitar dados específicos, não publicados, cedidos em suporte papel ou magnético já sob a forma de *output*.

---

<sup>11</sup> Enquadra-se neste âmbito o projecto de investigação *Padrões de vida: perfis e tendências na sociedade portuguesa contemporânea (Desenvolvimentos conceptuais e exploração analítica de microdados estatísticos)*, desenvolvido pelos autores do presente artigo. Trata-se de um projecto em curso no Centro de Investigação e Estudos de Sociologia (CIES), centro associado do Instituto Superior de Ciências do Trabalho e da Empresa (ISCTE), apoiado pela Fundação para a Ciência e Tecnologia (FCT), do Ministério da Ciência e da Tecnologia (MCT). A contribuição do INE para este projecto é também fundamental na constituição da equipa e na disponibilização das bases de microdados estatísticos.

As estatísticas divulgadas nestes três modelos são fornecidas, geralmente, sob a forma de tabelas de frequências ou cruzamentos simples de indicadores agregados. Indicadores esses construídos, muitas vezes, a partir de procedimentos estandardizados, os quais, sendo comuns aos diversos organismos nacionais e internacionais, apresentam a vantagem de permitir a comparação dos dados.

Não obstante, esta informação é, muitas vezes, insuficiente para o desenvolvimento das análises que os investigadores se propõem realizar. Nestes três modelos de disponibilização de dados é muito difícil, ou mesmo impossível, o desenvolvimento de um conjunto de operações muito importantes na actual investigação em ciências sociais. Em concreto, não é em geral possível proceder ao ensaio de categorizações e agregações alternativas, à construção e teste de novas variáveis derivadas, à realização de análises multivariadas de diversos tipos. Tudo isso precisando de ser conduzido, por um lado, em função de orientações teórico-conceituais muito complexas ou específicas e, por outro, em regime de vai-vem experimental entre sucessivas variantes operatórias e análises dos respectivos resultados.

Em Portugal, os condicionalismos legais respeitantes à necessidade de preservação do segredo estatístico impõem restrições rigorosas na disponibilização da informação sobre a forma de microdados (CSE, 1998). Isto, mesmo quando as pretensões de manuseamento dessa informação se circunscrevem à investigação científica. O acesso a este tipo de informação é uma possibilidade apenas disponível com a celebração do referido *Protocolo entre o Ministério da Ciência e da Tecnologia e o Instituto Nacional de Estatística*.

Neste protocolo reconhece-se a necessidade de disponibilizar mais e melhor informação para uso na investigação científica, ao mesmo tempo que se regulamenta a sua utilização. É prestada especial atenção à necessidade de protecção dos direitos individuais, sobretudo no que diz respeito ao anonimato e privacidade. Neste sentido, o protocolo sugere que o INE, a instituição detentora dos dados, proceda à anonimização prévia de todas as bases de dados com informação de natureza individual. Cabe, pelo seu lado, ao MCT a credenciação das instituições de investigação científica e desenvolvimento tecnológico, cujos investigadores pretendam aceder a este tipo de informação. Adicionalmente, há uma responsabilização individualizada dos investigadores que irão utilizar a informação. A cada um é exigida a assinatura de um termo de responsabilidade que o obriga a respeitar um conjunto de regras, entre as quais se destacam:

- utilizar os dados estatísticos apenas para os fins da investigação científica especificados num resumo do projecto previamente apresentado;
- não procurar identificar indivíduos ou instituições envolvidos nas bases de dados;
- não comercializar ou ceder os dados a terceiros;
- referenciar o INE em qualquer publicação ou comunicação baseada nos dados fornecidos;
- informar o INE de quaisquer problemas de natureza técnica, metodológica, de registo de dados ou de erros detectados nas bases de dados utilizadas.

### 3. MODELOS DE ANONIMIZAÇÃO

A obrigação de anonimização de bases de dados colectados por organismos públicos está inscrita nos regulamentos do Conselho Europeu (CE) relativos às estatísticas comunitárias. Como salienta o regulamento do CE n. 322/97, “qualquer utilização dessa informação deve assegurar a impossibilidade prática de identificação das unidades estatísticas de registo envolvidas”, como sejam indivíduos, famílias, empresas, etc. Reconhece-se, assim, a necessidade de anonimização, ou seja, a necessidade de “anulação de todos os meios que possam razoavelmente ser utilizados por uma terceira parte para identificar a dita unidade estatística”.

Existem presentemente diversos métodos disponíveis para prevenir o risco de identificação de uma unidade de registo que conste na base de microdados (Dale, 2001). Importa sublinhar, antes de mais, que os procedimentos de anonimização não são independentes das características dos microdados em causa. Por exemplo, as soluções de anonimização podem ser, em parte, diferentes consoante se trate de dados censitários ou de dados recolhidos através de processos de amostragem. Essas diferenças podem ter a ver também com outros aspectos, designadamente com o tipo de unidades de registo – se se trata de empresas ou de edifícios, de agregados familiares ou de indivíduos ou de outras entidades.

Quando estão em causa universos de grande dimensão e dados recolhidos por amostragem, a eliminação das variáveis de identificação das unidades de registo constitui com frequência, só por si, um procedimento eficaz de anonimização. Ainda assim, pode ser conveniente activar outros procedimentos de anonimização e, por maioria de razão, quando o número de casos é reduzido para certas categorias ou combinações de categorias.

Os modelos de anonimização, enquanto componentes de estratégias visando a disponibilização de bases de microdados para a investigação científica, podem ser agrupados como segue.

#### A) A CONSTRUÇÃO DE SAFE VARIABLES: AGREGAÇÃO DE CATEGORIAS

A construção de *safe variables* constitui uma das metodologias mais utilizadas. Esta consiste na selecção de variáveis contidas nas bases de dados originais – as variáveis que poderiam conduzir à identificação de unidades estatísticas –, as quais são sujeitas a um processo de *anonimização por agregação de categorias*. Este método pode ser utilizado quando não é possível assegurar a confidencialidade dos dados originais, por exemplo, devido ao nível de detalhe de uma determinada variável (unidade geográfica, profissional, etária, etc.).

A operacionalização destes procedimentos de anonimização permite soluções diversas consoante o tipo de utilização que se pretende dar aos dados. Na selecção das variáveis a agregar procura-se ter em conta as perspectivas de análise e os temas específicos de investigação. Ou seja, se, por exemplo, em algumas pesquisas pode ser vantajoso construir a anonimização com base numa forte agregação do âmbito geográfico, permitindo maiores especificações de variáveis educacionais ou profissionais, por

exemplo, noutras, onde os indicadores de caracterização espacial fina correspondam a vectores-chave da investigação, o balanço mais útil pode ir no sentido inverso, passando pela agregação de outros indicadores, nomeadamente níveis etários, grupos profissionais, etc.

#### B) A CONSTRUÇÃO DE SAFE SETTINGS: AGREGAÇÃO DE CASOS

Uma alternativa ao procedimento anterior é *agregar*, não as categorias de variáveis, mas *os casos em subconjuntos*, atribuindo a cada caso o valor de caracterização do subconjunto.

O acesso em linha aos microdados apoiado por um *safe setting* foi desenvolvido, designadamente no Luxemburgo, no âmbito do CEPS/INSTEAD.<sup>12</sup> Neste modelo, os investigadores munidos de uma *password*, acedem, também, a uma subamostra representativa da base de sondagem. As variáveis dos vários inquéritos são previamente harmonizadas para permitir a comparabilidade, o que constitui por si só um grande benefício para os investigadores (Förster, Helliesen e Kolberg, 1996; Christos, 2000; Dale, 2001).

No entanto, este método de trabalho, comparativamente à metodologia anterior, apresenta algumas desvantagens. Entre elas, o facto de o processo de agregação ser mais complexo e introduzir maior indeterminação na análise. Para além disso o nível de interacção com os dados é limitado, aumentando a barreira à construção e teste de novas variáveis.

#### C) A CONSTRUÇÃO DE SUBAMOSTRAS DAS BASES DE DADOS ORIGINAIS

A construção de subamostras dos dados originais, como forma de facilitar o acesso aos microdados em formato digital, é utilizada, por exemplo no Reino Unido, desde 1993, para as *estatísticas censitárias* (Dale, 2001). Parte-se do princípio de que, com a construção destas subamostras, fica assegurado o anonimato, com a impossibilidade de identificação dos dados. A informação assim disponibilizada é acessível tanto para fins de investigação como para o ensino. Cada universidade foi convidada a nomear um membro que assume a responsabilidade pelo uso dos dados no quadro da respectiva instituição. Adicionalmente, professores ou investigadores que queiram utilizar a informação estatística assinam um compromisso que os obriga pessoalmente a manterem a segurança dos dados.

#### D) A DISPONIBILIZAÇÃO DE MICRODADOS EM CENTROS AUTORIZADOS

Um outro modelo alternativo, também disponível, por exemplo no, já referido, CEPS/INSTEAD, conjuga o *safe setting* com a disponibilização das bases de dados em centros autorizados onde os investigadores podem permanecer durante determinado período para o desenvolvimento de análises específicas. O *safe setting* assume, neste caso, a forma de um cálculo laboratorial seguro, onde é também rigorosamente impedida a alteração dos registos originais. Este método faculta a oportunidade de tratamento de dados de todos os casos colectados e não apenas de uma subamostra, como acontecia no caso anterior. A implementação deste processo de anonimização é, no

<sup>12</sup> Site: <http://lisweb.ceps.lu/lis>

entanto, muito dispendiosa, implicando a existência de uma equipa permanente que dê apoio e supervisione os investigadores. Além disso, custos e inconvenientes são também impostos aos investigadores, que têm de se deslocar para trabalharem com todos os dados, podendo implicar que tenham de aí permanecer algum tempo até concluírem a pesquisa.

#### E) O ACESSO INDIRECTO AOS DADOS

Outro processo que não envolve microdados, mas que em muitos casos resolve as necessidades dos utilizadores, consiste no acesso indirecto aos dados através da elaboração de pedidos específicos, dirigidos à instituição produtora/gestora das bases de dados, de acordo com as análises que se pretendem conduzir. Este sistema permite solicitar frequências ou cruzamentos entre indicadores. Porém, não permite, por exemplo, a criação de novas variáveis ou a elaboração de testes estatísticos mais sofisticados, designadamente de estatística multivariada, nem possibilita experimentação indirecta de sucessivos tratamentos alternativos, o que o torna menos atractivo para alguns tipos de investigação.

#### F) O ACESSO ATRAVÉS DA INTERNET

Uma situação intermédia entre a referida anteriormente e os vários modelos de acesso directo aos microdados, consiste na disponibilização de bases de dados standardizadas na Internet. Este modelo possibilita a realização de apuramentos simples de frequências, a selecção de categorias de variáveis e/ou segmentos da população, bem como, pedidos de cruzamentos com duas ou mais variáveis. Para além de facultar a experimentação das bases de dados, tem a vantagem de integrar na informação disponibilizada todo um conjunto de variáveis derivadas de uso comum (como a condição perante o trabalho, diversas taxas de emprego, etc.), cuja construção implicaria o domínio de algoritmos específicos. A nível nacional, esta poderá ser a orientação dominante na disponibilização de informação estatística por parte das instituições produtoras de microdados.

---

## 4. MICRODADOS ESTATÍSTICOS E META-INFORMAÇÃO

---

No uso de microdados estatísticos, outra questão relevante, também ela respeitante às condições de disponibilização deste tipo de informação, tem a ver com a meta-informação associada a cada projecto do INE (veja-se a este respeito Fonseca, Valente e Lima, 1995). Esta refere-se à informação conceptual e metodológica sobre cada base de dados e sobre as operações estatísticas que a ela conduziram.

A utilização informada de bases de microdados estatísticos requer, nomeadamente, informação especificada sobre:

- a *metodologia do inquérito* (método de recolha, questionário, características da amostra, unidades estatísticas, método de tratamento de não-respostas, nomenclaturas);

- as *metodologias de estimação* (cálculo dos ponderadores e sua utilização); estas podem assentar em perspectivas diversas, atendendo a objectivos diferenciados: uma, como forma de corrigir pesos relativos para trabalhar a amostra com probabilidades corrigidas; outra, é a sua utilização como critério de extrapolação, quando se quer trabalhar com valores absolutos;
- as *variáveis e seus conceitos* (descritivos das variáveis, incluindo as variáveis derivadas, tabelas de codificação e condições de resposta ou filtros de selecção de casos);
- a *informação técnica* relativa às bases de dados (desenho de registo e *software* utilizado);
- os critérios de *interpretação e divulgação* (metodologia de cálculo dos erros de amostragem associados aos resultados e sua utilização).

---

## 5. VALIDAÇÃO E CONTROLO DA QUALIDADE: CONCEPÇÕES E PROCEDIMENTOS PARA A UTILIZAÇÃO DOS MICRODADOS

---

A utilização adequada de microdados não pode deixar de ser acompanhada de uma garantia de qualidade e validação da informação produzida. Neste sentido, torna-se crucial que se conheçam as metodologias e procedimentos técnicos de recolha e tratamento da informação, dos quais se podem destacar:

- em primeiro lugar, os *métodos de amostragem* e possibilidades de representatividade das amostras construídas;
- em segundo lugar, as *formas de cálculo dos ponderadores*;
- em terceiro lugar, os procedimentos de *validação* sistemática de recolha, tratamento e análise dos dados.

Em relação aos *métodos de amostragem*, o controlo dos processos, com a consciência dos limites e potencialidades associados a esses métodos, supõe que o investigador tenha informação, nomeadamente no que diz respeito à população seleccionada, à base de sondagem, à dimensão e métodos de selecção da amostra e critérios de estratificação que lhe estão associados, entre outros (Azorín e Sánchez-Crespo, 1986; Ghiglione e Matalon, 1993; Vicente, Reis e Ferrão, 1996).

Outra questão de relevo no uso das bases de microdados prende-se com o *cálculo dos ponderadores* e sua utilização. A opção por uma determinada metodologia pode estar associada a objectivos de utilização diversificados:

- na investigação em ciências sociais, a sua utilização tem muitas vezes como objectivo a correcção da proporcionalidade no cálculo das relações entre indicadores. Aqui os valores ponderados apresentam-se, regra geral, em termos relativos ou em médias;

- outro objectivo de utilização frequentemente operacionalizado pelo INE, pondera os dados da amostra para estimar os valores absolutos do universo. Neste caso, uma vez o ponderador activo, os dados apurados reflectem em termos absolutos as ordens de grandeza da população.

No que se refere à *validação da informação*, um dos aspectos a ter em conta prende-se com a eventual transposição da base de microdados de uma aplicação informática para outra, um processo que muitas vezes requer o apoio técnico do INE. Em seguida, quando da recepção dos dados, é de grande importância, como se referiu, a familiarização com a metainformação. O questionário e os desenhos de registos são conteúdos incontornáveis deste instrumento, nomeadamente para a validação da estrutura da base de dados e para o reconhecimento das variáveis.

Um outro procedimento de controlo no uso de microdados é a comparação entre as questões (e respectivas modalidades de resposta) que constam no instrumento de recolha de informação, e as variáveis presentes na base de dados recebida.

Para o utilizador externo à comunidade produtora dos dados, o controlo de qualidade e conseqüente validade dos resultados poderá passar, numa primeira fase e sempre que possível, pela comparação desses apuramentos com estatísticas já publicadas pelo organismo responsável pela sua recolha e difusão, e que tenham por base as mesmas definições das variáveis. Uma análise crítica da coerência e consistência dos apuramentos realizados, nomeadamente através de uma perspectiva longitudinal, poderá ainda consolidar essa validação.

A validação engloba ainda o conhecimento sobre mudanças de séries dos questionários e das respectivas aplicações (Fonseca, Lourenço e Lima, 1995). Tais modificações podem decorrer de melhoramentos na cobertura do inquérito em causa, correcções sazonais, aperfeiçoamentos metodológicos, etc. (Suominen, 2001).

---

## 6. CONDIÇÕES DE DIVULGAÇÃO DE ANÁLISES APOIADAS EM MICRODADOS

---

As exigências de utilização de microdados alargam-se também às condições de divulgação e difusão de análises apoiadas em informação estatística.

Os procedimentos associados à divulgação dos dados obrigam a que sejam cumpridos os requisitos acordados entre INE e investigadores, no sentido de garantir rigor tanto na informação como na sua referenciação. Nesta medida, é também importante que se estabeleçam alguns compromissos de divulgação de informação e análise estatística por parte dos utilizadores.

Um desses compromissos, que se pode designar como *ético*, prende-se com as garantias de anonimização que o investigador deve dar quando publica ou difunde informação deste tipo. Nesta perspectiva, o rigor que pauta a produção de conhecimento e respectiva divulgação está profundamente associado a preocupações cívicas,

no sentido da garantia dos direitos e das liberdades, sobretudo no que respeita à defesa da privacidade e da autonomia individuais.

Outro compromisso, não menos importante, é de carácter *técnico*. Este prende-se fundamentalmente com preocupações ligadas à qualidade dos dados, tais como o controlo da significância, a construção de novas variáveis e, na divulgação destas, a utilização de designações distintas das divulgadas pelo INE, para não se confundirem as respectivas definições.

---

## 7. HARMONIZAÇÃO E COMPARABILIDADE DOS DADOS

---

Uma outra questão que tem estado muito presente nas discussões da comunidade produtora de estatísticas é a que se refere à comparabilidade de informação entre países. Esta é concretizada, em grande parte, através da estandarização das metodologias de recolha de informação e/ou de variáveis entre os vários países membros da União Europeia, tendo necessariamente repercussões na utilização de microdados.

Através da harmonização procura-se conjugar estatísticas de fontes diversificadas, que têm como objectivo tornar os conceitos e os métodos utilizados claros e comparáveis entre si, no sentido de se apresentar um só quadro de informação estatística coerente. Tal procedimento permite, por um lado, que a informação possa ser lida cada vez mais numa perspectiva longitudinal e, por outro, que seja possível, de forma mais articulada, construir análises apoiadas em estatísticas que atravessem e, se possível, cruzem várias áreas, nomeadamente, no âmbito dos países da União Europeia.

A possibilidade de comparação constitui uma vantagem não só no âmbito das decisões económicas e políticas tomadas a nível europeu, mas também para um tipo de investigação que se pretenda situar numa escala europeia ou mundial (veja-se a este respeito Costa, Mauritti, Martins, Machado e Almeida, 2000: 9-46). Acedendo aos microdados os seus utilizadores podem, para além de reforçar a possibilidade de comparar informação construída segundo conceitos uniformizados, enriquecer as suas análises, dotando-as, de forma controlada, de especificidades de âmbito nacional, e testar os limites e a justeza dos conceitos harmonizados.

Quando o esforço de harmonização se centra essencialmente na recolha e compilação dos dados (estratégia *input*), a utilização de microdados permite controlar, de forma mais ponderada, constrangimentos que se prendem com as especificidades de jurisdição e legislação nacionais face à regulamentação da produção e uso estatísticos, bem como ao entendimento e cultura de informação que existe em cada país e Estado. As formas de ponderação e os níveis de significância, conhecendo-se as especificidades dos métodos de recolha de informação, podem aqui permitir, como referido, um leque de escolha ao utilizador que lhe possibilite um uso mais adequado às suas finalidades.

Os processos de harmonização, uma vez focalizados no tratamento de dados (estratégia *output*), estratégia de maior recorrência no âmbito das estatísticas europeias

(Jensen, 2000), colocam outros desafios aos utilizadores de microdados. Tais processos tendem a uniformizar conceitos, exigindo aos utilizadores a respectiva interpretação de forma contextualizada, apoiada tanto quanto possível no conhecimento da diversidade social e cultural, das estruturas e dinâmicas sociais próprias de cada país.

O uso de microdados permite, não se perdendo a perspectiva da comparabilidade, que se conheçam as agregações implícitas às variáveis uniformizadas e que se controle a qualidade imposta pela estandarização das variáveis. Através do acesso às bases de dados oficiais, tal comparabilidade pode ser ainda enriquecida através de especificações regionais ou nacionais, em associação ou complementarmente à identificação de perfis ou padrões construídos através da experimentação e realização de análises multivariadas, que permitem ainda assegurar a qualidade dos conceitos comparados.

As estratégias para a harmonização de dados implicam muitas vezes mudanças e interrupções dos inquéritos, o que cria constrangimentos a leituras longitudinais. Esta questão não deixa de ser relevante, também, quando se trata da utilização de bases de microdados, pois a mudança de variáveis e das suas formas de categorização torna as comparações entre as várias séries de inquéritos mais limitadas. No entanto, como referido, não se pode estancar os processos de mudança metodológica dos inquéritos e de definição de conceitos, pois, muitas vezes, vão no sentido do seu aperfeiçoamento ou exprimem dinâmicas de mudança na própria sociedade.

---

## 8. NOTAS FINAIS

---

A utilização de microdados estatísticos produzidos pelo INE, possibilidade aberta recentemente com o estabelecimento do protocolo entre o MCT e esse instituto, constitui um importante passo no desenvolvimento da investigação em ciências sociais em Portugal.

No centro da exploração deste tipo de informação está a possibilidade de combinar novas propostas teóricas com a exploração de novas alternativas de recurso à informação empírica. Esta consubstancia-se no facto de se poder construir e ensaiar modos alternativos de utilização de dados estatísticos. Entre estes, destaca-se a possibilidade de criação de novas variáveis, a oportunidade de realização de análises estatísticas multivariadas e, em parte através destas, a capacidade de construção e aperfeiçoamento de tipologias e modelos de análise. Todos estes processos podem ser ainda desenvolvidos e aprofundados através de uma perspectiva experimental no uso continuado das bases de dados.

Nem todas as situações de investigação exigem, ou mesmo aconselham, a utilização de microdados. Para muitas análises são suficientes as formas de acesso que já existiam anteriormente.

A divulgação de microdados supõe o assumir de um compromisso entre a disponibilização de mais e melhor informação e a garantia de respeito da salvaguarda de

direitos dos cidadãos, nomeadamente no que concerne à defesa de privacidade e autonomia individuais (Trivellato, 2000). Decorrem daqui alguns dos princípios básicos estabelecidos no âmbito do protocolo, como sejam a credenciação das instituições de investigação científica e tecnológica, a responsabilização dos seus investigadores e a anonimização das bases de dados por parte do INE.

Importa salientar que a utilização de microdados tem subjacente o domínio de alguns aspectos processuais de natureza técnica, transversais a todo o processo de produção, tratamento e difusão dos dados estatísticos. Nesta medida, conjugam-se aspectos cruciais ligados, entre outros, à metainformação, às formas de estimação dos resultados, à harmonização e comparabilidade.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ANDROVITSANEAS, CHRISTOS, "New technologies in statistics and requirements for central institutions: from a user's perspective", Innovation in Provision and Production of Statistics: The Importance of New Technologies, conferências do CEIES, Helsínquia, 20-21 de Janeiro de 2000 (policopiado).
- AZORÍN, FRANCISCO; e SANCHÉZ-CRESPO, José Luís, Métodos e Aplicacions del Amuestro, Madrid, Alianza Universidad Textos, 1986.
- COSTA, ANTÓNIO FIRMINO DA; MAURITTI, ROSÁRIO; MARTINS, SUSANA DA CRUZ; MACHADO, FERNANDO LUÍS; e ALMEIDA JOÃO FERREIRA DE, "Classes sociais na Europa", Sociologia, Problemas e Práticas, 34, 2000, pp. 9-46.
- CSE, Relatório de Actividades do Conselho Superior de Estatística, 1998, DOCT/351/CSE-2, 1998 (policopiado).
- DALE, ANGELA, "An academic perspective on the information society", Challenges for Community Statistics: A Users' Perspective for the Next Ten Years, conferências do CEIES Bruxelas, 26-27 de Abril de 2001 (policopiado).
- FONSECA, ISABEL DUARTE, VALENTE, ISABEL LOURENÇO, e LIMA, JOSÉ PEDRO, "A gestão da metainformação no sistema estatístico português", Sociologia, Problemas e Práticas, 17, 1995, pp. 153-168.
- FÖRSTER, MICHAEL; HELLIESEN, ANNE; e KOLBERG, JOHN-EIVIND, "LES, the new challenge: an introduction into the concepts, documentation and background statistics", Luxembourg Employment Study, Working Paper, 1, 1996, em <http://lisweb.ceps.lu/les/wps/leswp1.htm>.
- GHIGLIONE, RODOLPHE; e MATALON, BENJAMIM (1993), Inquérito: Teoria e Prática, Oeiras, Celta Editora.
- JENSEN, LOTHAR, "Prospects for European social statistics: from Maastricht to Lisbon", Sigma: The Bulletin of European Statistics, 3/2000, pp. 2-4.

LAMEL, JOACHIM, "Feeling the pulse of society", Sigma: The Bulletin of European Statistics, 2/2001, pp. 2-3.

PROTOCOLO ENTRE O MINISTÉRIO DA CIÊNCIA E DA TECNOLOGIA E O INSTITUTO NACIONAL DE ESTATÍSTICA PARA ACESSO A BASES DE DADOS PARA FINS DE INVESTIGAÇÃO CIENTÍFICA, 18 de Outubro de 1999.

SUOMINEM, RISTO, "Enlargement and Deepening", Challenges for Community Statistics: A Users' Perspective for the Next Ten Years, conferências do CEIES, Bruxelas, 26-27 de Abril de 2001 (policopiado).

TRIVELLATO, UGO, "Data Access versus privacy: an analytical user's perspective", Innovation in Provision and Production of Statistics: The Importance of New Technologies, conferências do CEIES, Helsínquia, 20-21 de Janeiro de 2000 (policopiado).

VICENTE, PAULA; REIS, ELIZABETH; FERRÃO, FÁTIMA, Sondagens: Amostragem Como Factor Decisivo, Lisboa, Edições Sílabo, 1996.

## Dualidades entre Análise de Covariância e Análise de Regressão com variáveis dummy

Autores:  
Patrícia Oom do Valle  
Efigénio Rebelo

**VOLUME II**

**2º QUADRIMESTRE DE 2002**

---

---

## DUALIDADES ENTRE ANÁLISE DE COVARIÂNCIA E ANÁLISE DE REGRESSÃO COM VARIÁVEIS DUMMY

---

---

---

---

## DUALITIES BETWEEN ANCOVA AND REGRESSION WITH DUMMY VARIABLES

---

---

Autores: Patrícia Oom do Valle

- Assistente da Faculdade de Economia da Universidade do Algarve; Área Científica de Métodos Quantitativos.

Efigénio Rebelo

- Professor Associado da Faculdade de Economia da Universidade do Algarve; Área Científica de Métodos Quantitativos.

### RESUMO:

- A análise de covariância pode ser entendida como um elo de ligação entre a análise de variância e a análise de regressão uma vez que um dos seus principais objectivos é avaliar o efeito de um ou mais factores explicativos de natureza nominal numa dada variável resposta uma vez removida a influência que um ou mais factores quantitativos podem também exercer nessa variável. Neste estudo demonstra-se que através da especificação de um modelo de regressão com regressores dummy é possível atingir de uma forma muito eficiente este objectivo particular da análise de covariância.

### PALAVRAS-CHAVE:

- *Regressão linear, Variáveis dummy, Análise de covariância.*

### ABSTRACT:

- The purpose of the current study is to show how the specification of a linear regression model with dummy variables and the performance of some statistical tests enable to achieve very efficiently the main purpose of analysis of covariance (ANCOVA), that is, to assess the influence of two or more nominal variables in a response variable, once controlled the potential effect of one or more quantitative factors. In this sense, ANCOVA may be understood as the link between analysis of variance (ANOVA) and linear regression.

### KEY-WORDS:

- *Linear regression, Dummy variables, ANCOVA.*

**VOLUME II**

**2º QUADRIMESTRE DE 2002**

## 1. INTRODUÇÃO

A análise de covariância é uma técnica estatística com alguma complexidade pela quantidade significativa de cálculos que envolve. Como facilmente se compreende, o nível de sofisticação deste método é crescente com o número de factores explicativos nominais cuja significância se pretende testar bem como com o número de variáveis quantitativas que importa controlar. Naturalmente, este problema dissipa-se quando a experiência em causa pode ser conduzida através de programas informáticos adequados. Uma vez que o problema de análise de covariância em que se incidirá neste estudo será desenvolvido de uma forma descritiva, sem recurso portanto a 'packages' de cálculo, apenas um factor nominal com  $k$  níveis e um factor quantitativo serão considerados como fontes de variação da variável resposta.

## 2. OBJECTIVOS E PRESSUPOSTOS DA ANÁLISE DE COVARIÂNCIA

Considere-se o seguinte exemplo de uma experiência em que a análise de covariância pode ser aplicada. Um investigador pretende determinar se o horário em que uma determinada actividade fabril é desempenhada tem um efeito significativo na produtividade do trabalho. Suponha-se que a entidade que conduzirá a experiência tem razões para acreditar que o número de anos de experiência profissional de cada trabalhador constitui outro factor que não pode ser desprezado na avaliação da sua produtividade. Assim, uma eventual diferença de produtividade média entre os dois horários pode dever-se não ao facto de horários diferentes propiciarem efectivamente níveis de produtividade distintos, mas ocorrer sim em resultado dos trabalhadores com mais experiência exercerem as suas funções predominantemente num dos horários. Está-se portanto perante uma situação em que é desejável avaliar a influência do factor nominal "horário de trabalho" (também designado por factor experimental) na variável resposta "número de unidades produzidas por hora" uma vez eliminado o efeito da variável quantitativa (também designada por covariável, variável de controle ou variável concomitante). A análise de covariância apresenta-se como uma técnica apropriada para atingir este propósito.

O desenrolar da experiência pressupõe a selecção de uma amostra de  $n$  indivíduos dos quais  $n_i$  ( $i = 1, 2$ ) são aleatoriamente colocados a exercer funções em cada um dos horários. No início da primeira sessão experimental, regista-se o número de anos de experiência profissional de cada trabalhador e, uma vez terminada a experiência, o número de unidades produzidas por cada indivíduo é contabilizado. A aplicação da análise de covariância permitirá verificar se existem de facto diferenças de produtividade provocadas pelo exercício da actividade produtiva em horários diferentes, qualquer que seja a experiência profissional dos trabalhadores.

A análise de covariância é conhecida também como uma técnica que permite a comparação de várias equações de regressão (Dunn, 1974). Embora, numa primeira

análise, esta comparação pareça não estar relacionada com o conceito introduzido acima, não é difícil demonstrar que as duas perspectivas convergem e podem mesmo ser entendidas como duas faces de uma problemática comum.

Considere-se o caso mais geral em que o factor experimental se encontra dividido em  $k$  categorias nominais. Dunn expressa o modelo de análise de covariância da seguinte forma:

$$y_{ij} = \alpha_i + \beta_i X_{ij} + e_{ij}; i = 1, 2, \dots, k; j = 1, 2, \dots, n_i; e_{ij} \sim \text{IIN}(0, \sigma^2) \quad (1)$$

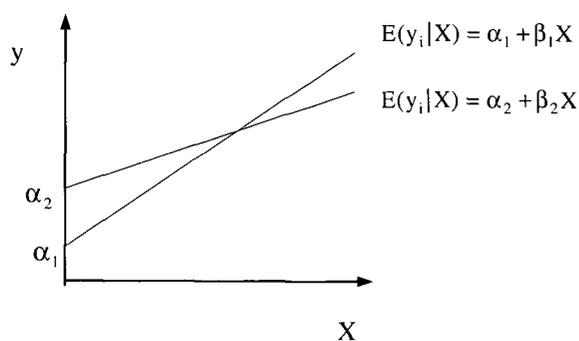
para deixar clara a existência de  $k$  regressões lineares distintas.

Os pressupostos subjacentes ao modelo (1) são as seguintes:

- i) Cada uma populações envolvidas na experiência segue uma distribuição normal;
- ii) As populações têm a mesma variância,  $\sigma^2$ ;
- iii) As amostras extraídas de cada população são aleatórias e independentes;
- iv) Existe uma relação linear entre a variável resposta e a covariável

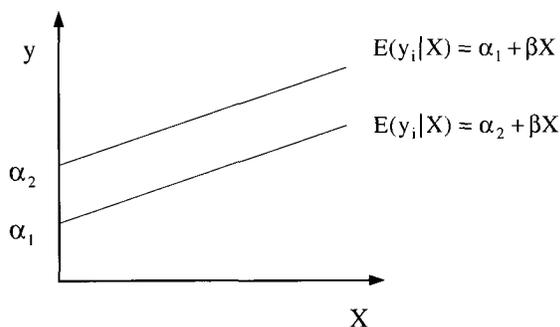
Admitindo que  $k = 2$ , três tipos de configurações possíveis, representadas nas figuras 1, 2 e 3, são ensaiáveis através da aplicação da análise de covariância. Note-se que as hipóteses subjacentes a cada construção gráfica estão indicadas na parte inferior de cada figura.

**Figura 1. Rectas de regressão com inclinação diferente**



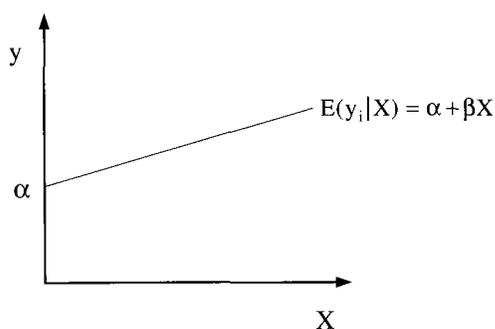
Hipótese nula subjacente:  $H_0^1: E(y_i | X) = \alpha_i + \beta_i X$   
( $i = 1, 2$ ; 1 = diurno e 2 = nocturno)

Figura 2. Rectas de regressão paralelas



Hipótese nula subjacente:  $H_0^2: E(y_i|X) = \alpha_i + \beta X$   
( $i = 1,2$ ; 1 = diurno e 2 = nocturno)

Figura 3. Rectas de regressão coincidentes



Hipótese nula subjacente:  $H_0^3: E(y_i|X) = \alpha + \beta X$   
( $i = 1,2$ ; 1 = diurno e 2 = nocturno)

Na figura 1, a distancia vertical entre as duas rectas de regressão é variável com os valores de X. Se a análise de covariância não rejeitar a hipótese subjacente à sua construção, não é possível concluir acerca da diferença entre as médias das populações envolvidas sem ter em atenção o valor da covariável. Em termos do exemplo descrito acima, isto significa que apenas se pode afirmar que “para elevados níveis de formação profissional a produtividade é superior quando a actividade é desenvolvida no horário diurno”.

A situação mais desejada é a que ilustra a figura 2. Se a análise de covariância confirmar o paralelismo entre as rectas de regressão, então sim pode concluir-se que “para qualquer nível de formação profissional, a produtividade é superior (em  $\alpha_1 - \alpha_2$ ) quando as várias funções são desempenhadas no período diurno”.

A figura 3 evidencia o caso em que as duas rectas de regressão coincidem. A análise de covariância permite também ensaiar a validade desta hipótese e, no contexto do exemplo, a sua não rejeição deve ser interpretada da seguinte forma: “não existem diferenças de produtividade entre os dois horários de trabalho qualquer que seja o número de anos de experiência profissional dos indivíduos envolvidos.

A atitude do investigador perante cada situação é necessariamente diferente. Se a hipótese  $H_0^1 : E(y_i | X) = \alpha_i + \beta_i X$  não for rejeitada, os parâmetros de cada uma das rectas de regressão deverão ser estimados em separado. Pelo contrário, se a hipótese  $H_0^2 : E(y_i | X) = \alpha_i + \beta X$  não for rejeitada, o investigador deverá proceder à estimação do parâmetro comum às duas rectas de regressão. Finalmente, se a hipótese  $H_0^3 : E(y_i | X) = \alpha + \beta X$  for confirmada, apenas uma equação de regressão deverá ser estimada. A análise de covariância permite testar cada uma destas hipóteses e, simultaneamente, estimar os coeficientes inclinação de cada uma das situações descritas.

### 3. O QUADRO DE ANÁLISE DE COVARIÂNCIA

Antes de se apresentar o quadro de análise de covariância, com base no qual os rácios necessários para testar estas hipóteses serão construídos, os seguintes cálculos preliminares, definidos para o caso mais geral em que a variável nominal apresenta  $k$  categorias, deverão ser realizados:

i) Soma de Quadrados para  $X$

$$SQX_1 = \sum_{j=1}^{n_1} (X_{1j} - \bar{X}_1)^2, SQX_2 = \sum_{j=1}^{n_2} (X_{2j} - \bar{X}_2)^2, \dots, SQX_k = \sum_{j=1}^{n_k} (X_{kj} - \bar{X}_k)^2 \quad (2)$$

$$SQX_T = SQX_1 + SQX_2 + \dots + SQX_k \quad (3)$$

$$SQX_{\text{entre}} = \sum_{i=1}^k \sum_{j=1}^{n_j} (\bar{X}_i - \bar{X})^2, SQX_{\text{dentro}} = \sum_{i=1}^k \sum_{j=1}^{n_j} (X_{ij} - \bar{X}_i)^2 \quad (4)$$

$$SQX_{\text{Total}} = SQX_{\text{entre}} + SQX_{\text{dentro}} \quad (5)$$

ii) Soma de Quadrados para  $y$

$$SQY_1 = \sum_{j=1}^{n_1} (y_{1j} - \bar{y}_1)^2, SQY_2 = \sum_{j=1}^{n_2} (y_{2j} - \bar{y}_2)^2, \dots, SQY_k = \sum_{j=1}^{n_k} (y_{kj} - \bar{y}_k)^2 \quad (6)$$

$$SQY_T = SQY_1 + SQY_2 + \dots + SQY_k \quad (7)$$

$$SQY_{entre} = \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y})^2, \quad SQY_{dentro} = \sum_{i=1}^k \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 \quad (8)$$

$$SQY_{Total} = SQY_{entre} + SQY_{dentro} \quad (9)$$

iii) Soma de Produtos para X e y

$$SPXY_1 = \sum_{j=1}^{n_1} (X_{1j} - \bar{X}_1)(y_{1j} - \bar{y}_1), \dots, \quad SPXY_k = \sum_{j=1}^{n_k} (X_{kj} - \bar{X}_k)(y_{kj} - \bar{y}_k) \quad (10)$$

$$SPXY_T = SPXY_1 + SPXY_2 + \dots + SPXY_k \quad (11)$$

$$SPXY_{entre} = \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{X}_i - \bar{X})(\bar{y}_i - \bar{y}), \quad SPXY_{dentro} = \sum_{i=1}^k \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)(y_{ij} - \bar{y}_i) \quad (12)$$

$$SPXY_{Total} = SPXY_{entre} + SPXY_{dentro} \quad (13)$$

O quadro correspondente à análise de covariância pode agora ser apresentado.

**Quadro 1. Quadro da Análise de Covariância**

FONTE DE VARIACÃO	SOMA DE QUADRADOS X	SOMA DE QUADRADOS Y	SOMA DE PRODUTOS XY	GRAUS DE LIBERDADE	COEFICIENTE DE REGRESSÃO	SOMA DE QUADRADOS RESÍDUOS	GRAUS DE LIBERDADE	QUADRADOS MÉDIOS
(1)	(2)	(3)	(4)	(5)	(6)=(4)/(2)	(7)=(3)-(6) <sup>2</sup> × (2)	(8)	(9)=(7)/(8)
Resíduo 1	SQX <sub>1</sub>	SQY <sub>1</sub>	SPXY <sub>1</sub>	n <sub>1</sub> - 1	$\hat{b}_1$	S.Q.R. <sub>1</sub>	n <sub>1</sub> - 2	
Resíduo 2	SQX <sub>2</sub>	SQY <sub>2</sub>	SPXY <sub>2</sub>	n <sub>2</sub> - 1	$\hat{b}_2$	S.Q.R. <sub>2</sub>	n <sub>2</sub> - 2	
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	
Resíduo k	SQX <sub>k</sub>	SQY <sub>k</sub>	SPXY <sub>k</sub>	n <sub>k</sub> - 1	$\hat{b}_k$	S.Q.R. <sub>k</sub>	n <sub>k</sub> - 2	
Total	SQX <sub>T</sub>	SQY <sub>T</sub>	SPXY <sub>T</sub>	N - k		S.Q.R. <sub>T</sub>	N - 2k	QMR <sub>T</sub>
Factor Explicativo	SQX <sub>entre</sub>	SQY <sub>entre</sub>	SPXY <sub>entre</sub>	k - 1		S.Q.R. <sub>entre</sub>	k - 1	
Resíduo	SQX <sub>dentro</sub>	SQY <sub>dentro</sub>	SPXY <sub>dentro</sub>	N - k	$\hat{b}$	S.Q.R. <sub>dentro</sub>	N - k - 1	QMR <sub>dentro</sub>
Total	SQX <sub>Total</sub>	SQY <sub>Total</sub>	SPXY <sub>Total</sub>	N - 1	$\hat{b}_T$	S.Q.R. <sub>Total</sub>	N	QMR <sub>entre</sub>

Descreve-se seguidamente os testes estatísticos que permitem ensaiar, no âmbito da ANCOVA, cada uma das três hipóteses acima referidas.

Assim, a estatística F que possibilita testar se as rectas de regressão têm a mesma inclinação é a seguinte

Assim, a estatística F que possibilita testar se as rectas de regressão têm a mesma inclinação é a seguinte

$$F_1 = \frac{(S.Q.R._{dentro} - S.Q.R._{t}) / \sum_{i=1}^k n_i - 2k}{S.Q.R._{t} / k - 1} \sim F \left[ k - 1, \sum_{i=1}^k n_i - 2k \right] \quad \left| \quad H_0^2 \right. \quad (14)$$

em que  $S.Q.R._{dentro}$  é o somatório do quadrado dos resíduos (S.Q.R.) que resulta da estimação de k regressões lineares paralelas e  $S.Q.R._{t}$  o S.Q.R. proveniente da estimação de k regressões lineares com inclinação diferente. Note que, neste caso, a hipótese nula a ensaiar é  $H_0^2 : E(y_i | X) = \alpha_i + \beta X$  e a hipótese alternativa é  $H_0^1 : E(y_i | X) = \alpha_i + \beta_i X$ . A regra de decisão para o teste é:

$$\text{Rejeitar } H_0^2 \text{ se } F > F_\alpha \left[ k - 1, \sum_{i=1}^k n_i - 2k \right]$$

$$\text{Não rejeitar } H_0^2 \text{ se } F \leq F_\alpha \left[ k - 1, \sum_{i=1}^k n_i - 2k \right].$$

Se este teste rejeitar  $H_0^2$  o investigador deverá estimar k equações de regressão com termos independentes e inclinações distintas recorrendo à seguinte fórmula

$$y_i = \bar{y}_i + \hat{b}_i (X - \bar{X}_i) \quad i = 1, 2, \dots, k \quad (15)$$

onde

$$\hat{b}_i = S.P.XY_i / S.Q.X_i. \quad (16)$$

Perante a não rejeição de  $H_0^2$ , importa testar se existem diferenças significativas entre os termos independentes das rectas de regressão ou, dito de outra forma, se estes coincidem ou não. A hipótese nula a testar é  $H_0^3 : E(y_i | X) = \alpha + \beta X$  e a hipótese alternativa  $H_0^2 : E(y_i | X) = \alpha_i + \beta X$ . A estatística com base na qual será conduzido este teste é

$$F_2 = \frac{(S.Q.R._{Total} - S.Q.R._{dentro}) / \sum_{i=1}^k n_i - k - 1}{S.Q.R._{dentro} / k - 1} \sim F \left[ k - 1, \sum_{i=1}^k n_i - k - 1 \right] \quad \left| \quad H_0^3 \right. \quad (17)$$

onde  $S.Q.R._{Total}$  é o S.Q.R. decorrente da estimação de uma única equação de regressão com todas as observações. A regra de decisão para o teste é:

$$\text{Rejeitar } H_0^3 \text{ se } F > F_\alpha \left[ k - 1, \sum_{i=1}^k n_i - k - 1 \right]$$

$$\text{Não rejeitar } H_0^3 \text{ se } F \leq F_\alpha \left[ k - 1, \sum_{i=1}^k n_i - k - 1 \right].$$

Se a hipótese  $H_0^3$  não for rejeitada, as  $k$  rectas de regressão coincidem e todas as observações das  $k$  amostras deverão ser combinadas para estimar uma recta de regressão única. A equação com base na qual esta recta será estimada é

$$y'' = \bar{y} + \hat{b}_T (X - \bar{X}) \quad (18)$$

em que

$$\hat{b}_T = SPXY_{(Total)} / SQX_{(Total)} \quad (19)$$

Mediante a rejeição de  $H_0^3$ , é possível dar uma resposta afirmativa à questão colocada no âmbito do exemplo apresentado, isto é, pode concluir-se que diferentes horários de trabalho propiciam diferentes níveis de produtividade qualquer que seja o número de anos de experiência profissional dos trabalhadores. Nesta situação, a equação que permitirá determinar cada uma das rectas de regressão paralelas é dada por:

$$y_i''' = \bar{y}_i + \hat{b} (X - \bar{X}_i) \quad i = 1, 2, \dots, k \quad (20)$$

com

$$\hat{b} = SPXY_{dentro} / SQX_{dentro} \quad (21)$$

Uma vez introduzidos os objectivos fundamentais da análise de covariância, bem como os procedimentos necessários à sua prossecução, importa revelar a relação estreita entre esta técnica estatística e a análise de regressão com uma variáveis explicativas nominais e quantitativas. Mais especificamente, mostrar-se-á a equivalência das estatísticas que permitem ensaiar as três hipóteses referidas no âmbito das duas técnicas e também que os modelos de regressão a ajustar em cada caso são similares.

---

#### 4. O MODELO DE ANÁLISE DE COVARIÂNCIA COMO UM MODELO DE REGRESSÃO NA PRESENÇA DE REGRESSORES DUMMY

---

A especificação mais geral do modelo de regressão linear com regressores dummy, com base no qual serão efectuados os testes de hipóteses acima apresentados no contexto da ANCOVA, é a seguinte:

$$y_{ij} = \delta_0 + \delta_1 H_{1j} + \beta_1 X_j + \beta_2 (H_{1j} X_j) + u_{ij};$$

$$j = 1, 2, \dots, n_i; \quad i = 1, 2; \quad u_{ij} \sim \text{IIN}(0, \sigma^2)$$
(22)

em que  $H_1$  é o regressor dummy representativo da variável nominal “horário de trabalho” e  $X$  a variável quantitativa “número de anos de experiência profissional”.  $H_1$  assume o valor 1 para os indivíduos que trabalham no horário diurno. Se designarmos  $H_1 X$  por  $Z$ , (22) é equivalente a:

$$y_{ij} = \delta_0 + \delta_1 H_{1j} + \beta_1 X_j + \beta_2 Z_j + u_{ij}; \quad j = 1, 2, \dots, n_i; \quad i = 1, 2; \quad u_{ij} \sim \text{IIN}(0, \sigma^2).$$
(23)

A finalidade desta secção é apresentar as expressões dos estimadores dos mínimos quadrados ordinários (E.M.Q.O.) dos parâmetros deste modelo. Assim, demonstra-se que os E.M.Q.O. para o *coeficiente da variável dummy* e *termo independente* são dados por<sup>13</sup>

$$\hat{\delta}_0 = \bar{y}_2 - \bar{X}_2 \hat{\beta}_1 - \bar{Z}_2 \hat{\beta}_2$$
(24)

$$\hat{\delta}_1 = (\bar{y}_1 - \bar{y}_2) - (\bar{X}_1 - \bar{X}_2) \hat{\beta}_1 - (\bar{Z}_1 - \bar{Z}_2) \hat{\beta}_2$$
(25)

ou, tendo em atenção que  $\bar{Z}_2 = 0$  e  $\bar{Z}_1 = \bar{X}_1$ ,

---

<sup>13</sup> Ver demonstração em Valle (1997), pp. 137-140.

$$\hat{\delta}_0 = \bar{y}_2 - \bar{X}_2 \hat{\beta}_1 \quad (26)$$

$$\hat{\delta}_1 = (\bar{y}_1 - \bar{y}_2) - (\bar{X}_1 - \bar{X}_2) \hat{\beta}_1 - \bar{X}_1 \hat{\beta}_2 \quad (27)$$

em que  $\bar{y}_i$  representa a produtividade média de um indivíduo que trabalhe no horário  $i$  ( $i=1,2$ ) e  $\bar{X}_i$  é a média amostral do número de anos de experiência profissional para cada grupo de observações ( $i=1,2$ )

Por sua vez, demonstra-se que os E.M.Q.O. para os *coeficientes das variáveis quantitativas* são dados pela fórmula<sup>14</sup>

$$\hat{\beta} = \left( \hat{u}_1 \hat{u}_1 \right)^{-1} \hat{u}_1 \hat{u}_2$$

em que  $\hat{u}_1$  representa o vector de resíduos da regressão de  $X$  nas variáveis dummy e variável associada ao termo independente e  $\hat{u}_2$  o vector dos resíduos da regressão de  $y$  também nessas variáveis. Da concretização do conteúdo de cada um dos vectores, é fácil mostrar que

$$\hat{u}_1 = M_w X = \begin{bmatrix} X_{11} - \bar{X}_1 & X_{11} - \bar{X}_1 \\ X_{12} - \bar{X}_1 & X_{12} - \bar{X}_1 \\ \vdots & \vdots \\ X_{1n_1} - \bar{X}_1 & X_{1n_1} - \bar{X}_1 \\ X_{21} - \bar{X}_2 & 0 \\ X_{22} - \bar{X}_2 & 0 \\ \vdots & \vdots \\ X_{2n_2} - \bar{X}_2 & 0 \end{bmatrix} \quad \hat{u}_2 = M_w y = \begin{bmatrix} y_{11} - \bar{y}_1 \\ y_{12} - \bar{y}_1 \\ \vdots \\ y_{1n_1} - \bar{y}_1 \\ y_{21} - \bar{y}_2 \\ y_{22} - \bar{y}_2 \\ \vdots \\ y_{2n_2} - \bar{y}_2 \end{bmatrix}$$

sendo  $M_w = I_n - W(W'W)^{-1}W'$  e  $W$  uma matriz constituída pela variável associada ao termo independente do modelo (23) e por uma variável dummy, isto é:

<sup>14</sup> Ver demonstração em Valle (1997), pp. 134-136.

$$W = \begin{bmatrix} 1 & 1 \\ \vdots & 1 \\ \vdots & \vdots \\ \vdots & 1 \\ \vdots & 0 \\ \vdots & 0 \\ \vdots & \vdots \\ \vdots & \vdots \\ 1 & 0 \end{bmatrix}.$$

Note que  $X$  representa o vector das observações relativas ao regressor quantitativo e  $y$  o vector das observações da variável dependente, ou seja:

$$X = \begin{bmatrix} X_{11} \\ X_{12} \\ \vdots \\ X_{1n_1} \\ X_{21} \\ X_{21} \\ \vdots \\ X_{2n_2} \end{bmatrix} \quad y = \begin{bmatrix} y_{11} \\ y_{12} \\ \vdots \\ y_{1n_1} \\ y_{21} \\ y_{21} \\ \vdots \\ y_{2n_2} \end{bmatrix}.$$

Nos vectores  $X$  e  $y$  as primeiras  $n_1$  observações dizem respeito aos trabalhadores que operam no horário diurno e as restantes  $n_2$  observações aos trabalhadores afectados ao horário nocturno.

A concretização das operações indicadas conduz a:

$$\left( \begin{matrix} \hat{u}_1 & \hat{u}_2 \end{matrix} \right)^{-1} = \begin{bmatrix} 1/\sum_{j=1}^{n_2} (X_{2j} - \bar{X}_2)^2 & -1/\sum_{j=1}^{n_2} (X_{2j} - \bar{X}_2)^2 \\ -1/\sum_{j=1}^{n_2} (X_{2j} - \bar{X}_2)^2 & 1/\sum_{j=1}^{n_1} (X_{1j} - \bar{X}_1)^2 + 1/\sum_{j=1}^{n_2} (X_{2j} - \bar{X}_2)^2 \end{bmatrix}$$

$$\hat{u}_1 \hat{u}_2 = \begin{bmatrix} \sum_{j=1}^{n_1} (X_{1j} - \bar{X}_1)(y_{1j} - \bar{y}_1) + \sum_{j=1}^{n_2} (X_{2j} - \bar{X}_2)(y_{2j} - \bar{y}_2) \\ \sum_{j=1}^{n_1} (X_{1j} - \bar{X}_1)(y_{1j} - \bar{y}_1) \end{bmatrix}$$

e, finalmente,

$$\hat{\beta}_1 = \frac{\sum_{j=1}^{n_2} (X_{2j} - \bar{X}_2)(y_{2j} - \bar{y}_2)}{\sum_{j=1}^{n_2} (X_{2j} - \bar{X}_2)^2} \quad (28)$$

$$\hat{\beta}_2 = \left[ \frac{\sum_{j=1}^{n_1} (X_{1j} - \bar{X}_1)(y_{1j} - \bar{y}_1)}{\sum_{j=1}^{n_1} (X_{1j} - \bar{X}_1)^2} - \frac{\sum_{j=1}^{n_2} (X_{2j} - \bar{X}_2)(y_{2j} - \bar{y}_2)}{\sum_{j=1}^{n_2} (X_{2j} - \bar{X}_2)^2} \right] \quad (29)$$

Outro aspecto a salientar, é que os E.M.Q.O. para os coeficientes das variáveis quantitativas são também determináveis ajustando duas equações de regressão separadas: a primeira com as observações que dizem respeito aos trabalhadores que desempenham as suas funções no horário diurno e a segunda com as restantes observações. Se se designar por  $\hat{b}_1$  e  $\hat{b}_2$  o estimador do coeficiente da variável quantitativa de cada um desses modelos, é possível demonstrar que:

$$\hat{\beta}_1 = \hat{b}_2 \quad (30)$$

$$\hat{\beta}_1 + \hat{\beta}_2 = \hat{b}_1 \quad (31)$$

Assim, as expressões para os E.M.Q.O. apresentadas em (26) e (27) são equivalentes a:

$$\hat{\delta}_0 = \bar{y}_2 - \bar{X}_2 \hat{b}_2 \quad (32)$$

$$\begin{aligned} \hat{\delta}_1 &= (\bar{y}_1 - \bar{y}_2) - (\bar{X}_1 - \bar{X}_2) \hat{b}_2 - \bar{X}_1 (\hat{b}_1 - \hat{b}_2) \\ &= (\bar{y}_1 - \bar{y}_2) + \bar{X}_2 \hat{b}_2 - \bar{X}_1 \hat{b}_1 \end{aligned} \quad (33)$$

## 5. TESTE AO PARALELISMO DAS RECTAS DE REGRESSÃO

O teste ao paralelismo das rectas de regressão, descrito no contexto da análise de covariância, encontra na análise de regressão com regressores dummy e regressores quantitativos a sua contrapartida no teste de significância individual ao coeficiente  $\beta_2$  do modelo (22). De facto, se este ensaio de hipótese não rejeitar,  $H_0^2: \beta_2 = 0$  concluí-se, com um nível de significância de  $\alpha \times 100\%$ , que todos os indivíduos reagem de forma uniforme a uma variação unitária da variável quantitativa, o que significa que as rectas de regressão representativas de cada um dos grupos de trabalhadores têm

a mesma inclinação ( $\beta_1$ , neste caso). Para ensaiar esta hipótese no contexto da análise de regressão, a seguinte estatística pode ser utilizada

$$F = \frac{\left( \hat{\hat{u}}_r \hat{\hat{u}}_r - \hat{\hat{u}} \hat{\hat{u}} \right) / g}{\hat{\hat{u}} \hat{\hat{u}} / (n-r)} \sim F_{(g, n-r)} \quad \left| \quad H_0^2 \right. \quad (34)$$

onde  $\hat{\hat{u}} \hat{\hat{u}}$  é o S.Q.R. do modelo (23),  $\hat{\hat{u}}_r \hat{\hat{u}}_r$  o S.Q.R. dessa equação de regressão sob a restrição  $\beta_2 = 0$ ,  $g = 1$  e  $n - r = n_1 + n_2 - 4$ , sendo  $r$  o número de variáveis independentes do modelo (23).

Demonstra-se seguidamente que a estatística  $F_1$  que permite ensaiar esta hipótese no contexto da ANCOVA (fórmula (14)) é equivalente ao rácio  $F$  definido em (34). Para determinar cada um dos S.Q.R. envolvidas na estatística (34), é necessário recorrer às igualdades (30) a (33). Assim,

$$\begin{aligned} \hat{u} &= y - \left( \hat{\delta}_0 + \hat{\delta}_1 H_1 + \hat{\beta}_1 X + \hat{\beta}_2 Z \right) = \\ &= \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} 1_1 & 1_1 & X_1 & X_1 \\ 1_2 & 0_2 & X_2 & 0_2 \end{bmatrix} \begin{bmatrix} \bar{y}_2 - \bar{X}_2 \hat{b}_2 \\ (\bar{y}_1 - \bar{y}_2) + \bar{X}_2 \hat{b}_2 - \bar{X}_1 \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_1 - \hat{b}_2 \end{bmatrix} = \\ &= \begin{bmatrix} \left( y_1 - \bar{y}_1 \right) - \left( X_1 - \bar{X}_1 \right) \hat{b}_1 \\ \left( y_2 - \bar{y}_2 \right) - \left( X_2 - \bar{X}_2 \right) \hat{b}_2 \end{bmatrix} \quad (35) \end{aligned}$$

onde  $1_i$  é um vector coluna unitário com  $n_i$  elementos ( $i = 1, 2$ ),  $0_2$  um vector nulo com  $n_2$  elementos,  $y_1$  o vector das primeiras  $n_1$  observações da variável dependente e  $y_2$  o vector das últimas  $n_2$  dessa variável. Por sua vez,  $X_1$  é um vector com  $n_1$  observações da variável  $X$  relativamente às quais o regressor dummy  $H_1$  assume o valor 1 e  $X_2$  é o vector com as restantes  $n_2$  observações da variável  $X$ .  $\bar{X}_i$  e  $\bar{y}_i$  ( $i = 1, 2$ ) representam os vectores que contêm as médias amostrais da variável quantitativa e da variável dependente para cada um destes grupos de observações.

O resíduo associado a cada observação define-se portanto como

$$\hat{u}_{ij} = (y_{ij} - \bar{y}_i) - (X_{ij} - \bar{X}_i) \hat{b}_i \quad i = 1, 2; k = 1, 2, \dots, n_i \quad (36)$$

donde resulta:

$$\hat{u}' \hat{u} = \sum_{j=1}^{n_1} (y_{1j} - \bar{y}_1)^2 + \sum_{j=1}^{n_2} (y_{2j} - \bar{y}_2)^2 - \hat{b}_1^2 \sum_{j=1}^{n_1} (X_{1j} - \bar{X}_1)^2 - \hat{b}_2^2 \sum_{j=1}^{n_2} (X_{2j} - \bar{X}_2)^2. \quad (37)$$

Uma vez encontrada a expressão de  $\hat{u}' \hat{u}$ , o passo seguinte consiste em definir matematicamente  $\hat{u}_r \hat{u}_r$ . Assim, da imposição da restrição  $\beta_2 = 0$  no modelo (23) vem

$$y_{ij} = \delta_0 + \delta_i H_{ij} + \beta X_{ij} + u_{ij}; j = 1, 2, \dots, n_i; i = 1, 2; u_{ij} \sim \text{IIN}(0, \sigma^2) \quad (38)$$

e, por um processo demonstrativo semelhante ao usado para determinar os E.M.Q.O. para os parâmetros do modelo (23)<sup>15</sup>, os E.M.Q.O. para o coeficiente da variável dummy, do termo independente e do coeficiente da variável quantitativa do modelo (38) são dados por:

$$\hat{\delta}_0 = \bar{y}_2 - \bar{X}_2 \hat{\beta} \quad (39)$$

$$\hat{\delta}_i = (\bar{y}_1 - \bar{y}_2) - (\bar{X}_1 - \bar{X}_2) \hat{\beta}. \quad (40)$$

$$\hat{\beta} = \frac{\sum_{i=1}^2 \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)(y_{ij} - \bar{y}_i)}{\sum_{i=1}^2 \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2}. \quad (41)$$

Finalmente, para determinar  $\hat{u}_r \hat{u}_r$ , recorre-se ao processo utilizado antes para encontrar  $\hat{u}' \hat{u}$ , isto é,

$$\hat{u}_r = y - \left( \hat{\delta}_0 + \hat{\delta}_i H_i + \hat{\beta} X \right) = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} 1_1 & 1_1 & X_1 \\ 1_2 & 0_2 & X_2 \end{bmatrix} \begin{bmatrix} \bar{y}_2 - \bar{X}_2 \hat{\beta} \\ (\bar{y}_1 - \bar{y}_2) - (\bar{X}_1 - \bar{X}_2) \hat{\beta} \\ \hat{\beta} \end{bmatrix} =$$

<sup>15</sup> Ver as igualdades (24), (25), (28) e (29).

$$= \begin{bmatrix} \left( y_1 - \bar{y}_1 \right) - \left( X_1 - \bar{X}_1 \right) \hat{\beta} \\ \left( y_2 - \bar{y}_2 \right) - \left( X_2 - \bar{X}_2 \right) \hat{\beta} \end{bmatrix} \quad (42)$$

e, em consequência,

$$\hat{u}_r' \hat{u}_r = \sum_{i=1}^2 \sum_{j=1}^{n_i} \left( y_{ij} - \bar{y}_i \right)^2 - \hat{\beta}'^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} \left( X_{ij} - \bar{X}_i \right)^2 \quad (43)$$

Observe que (37) não é mais do que o S.Q.R.<sub>r</sub> do rácio  $F_1$  apresentado em (14) no âmbito da análise de covariância, para o caso em que  $k=2$ . Similarmente, (43) traduz precisamente a mesma fórmula que define S.Q.R.<sub>dentro</sub> neste mesmo rácio, também para  $k=2$ . Portanto, fica demonstrado que os rácios (14) e (34) conduzem exactamente ao mesmo resultado.

Referiu-se na apresentação da análise de covariância, que no caso do teste F rejeitar  $H_0^2$ , a fórmula (15) deveria ser utilizada para estimar  $k$  equações de regressão distintas no âmbito da ANCOVA. Estas equações de regressão são igualmente determináveis através de um modelo de regressão do tipo (22), bastando para isso atribuir o valor adequado ao regressor dummy. Atendendo mais uma vez ao exemplo, para encontrar o modelo de regressão ajustado apenas com as observações relativas aos trabalhadores do horário diurno, é suficiente atribuir o valor 1 ao regressor dummy  $H_1$ , situação em que a estrutura estimada do modelo (22) pode reescrever-se como:

$$\hat{y}_{1j} = \left( \hat{\delta}_0 + \hat{\delta}_1 \right) + \left( \hat{\beta}_1 + \hat{\beta}_2 \right) X_j \quad j=1,2,\dots,n_1 \quad (44)$$

Do mesmo modo, para  $H_{1i} = 0$ , obtém-se o modelo de regressão ajustado para os trabalhadores do horário nocturno

$$\hat{y}_{2j} = \hat{\delta}_0 + \hat{\beta}_1 X_j \quad j=1,2,\dots,n_2 \quad (45)$$

Apesar de ligeiramente diferentes em aparência, (44) identifica-se com (15) para o caso em que  $i=1$  e (45) coincide com (15) para  $i=2$ . De facto, tendo em atenção (31), (32) e (33), o modelo (44) transforma-se em

$$\begin{aligned} \hat{y}_{1j} &= \left( \bar{y}_2 - \bar{X}_2 \hat{b}_2 \right) + \left( \bar{y}_1 - \bar{y}_2 \right) + \bar{X}_2 \hat{b}_2 - \bar{X}_1 \hat{b}_1 + \hat{b}_1 X_j = \\ &= \bar{y}_1 - \bar{X}_1 \hat{b}_1 + \hat{b}_1 X_j \end{aligned}$$

$$= \bar{y}_1 + \hat{b}_1 (X_j - \bar{X}_1) \quad (46)$$

e, substituindo (30) e (32) em (45) vem

$$\begin{aligned} \hat{y}_{2j} &= \bar{y}_2 - \bar{X}_2 \hat{b}_2 + \hat{b}_2 X_j \\ &= \bar{y}_2 + \hat{b}_2 (X_j - \bar{X}_2). \end{aligned} \quad (47)$$

Fica assim demonstrada a equivalência entre o modelo (15) – definido no contexto da ANCOVA – e o modelo (22). Como se provou, no âmbito da análise de regressão, este último modelo acaba por reduzir-se a (46) e a (47), tornando clara a correspondência entre as duas técnicas.

## 6. TESTE À COINCIDÊNCIA DAS RECTAS DE REGRESSÃO

Perante a não rejeição da hipótese  $H_0^2 : \beta_2 = 0$ , o investigador deverá eliminar a variável correspondente ( $H_1X$ ) do modelo e voltar a estimar a equação de regressão resultante. Com base neste novo modelo (que se consubstancia em (38)), a segunda hipótese formulada na análise de covariância pode facilmente ser ensaiada através da realização de um teste de significância individual ao coeficiente do regressor dummy,  $\delta_i$ . Na verdade, se a hipótese  $H_0^3 : \delta_i = 0$  for rejeitada, as rectas de regressão relativas às duas populações de indivíduos são paralelas não coincidentes e a diferença entre os seus termos independentes indica a diferença de produtividade entre indivíduos que trabalhem em horários distintos, qualquer que seja o nível de experiência profissional que possuam.

Assim, para demonstrar que o valor do rácio  $F_2$  em (17) coincide com o valor da estatística necessária para testar  $H_0^3 : \delta_i = 0$  no modelo de regressão (38), a estatística (34) será novamente utilizada. Neste caso particular,  $\hat{u}_i$  é dada pela igualdade (43),  $\hat{u}_i$  calcula-se reparametrizando o modelo (38) através da imposição da restrição  $\delta_i = 0$ . A equação de regressão assim obtida

$$y_{ij} = \delta_0 + \beta_r X_{ij} + u_{ij} ; i = 1, 2, \dots, j = 1, 2, \dots, n_j \quad (48)$$

produz os seguintes E.M.Q.O.

$$\hat{\delta}_0 = \bar{y} - \bar{X} \hat{\beta}_r \quad (49)$$

$$\hat{\beta}_T = \frac{\sum_{i=1}^2 \sum_{j=1}^{n_i} (X_{ij} - \bar{X})(y_{ij} - \bar{y})}{\sum_{i=1}^2 \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2} \quad (50)$$

e o resíduo associado a cada observação

$$\hat{u}_{ij} = y_{ij} - \left( \hat{\delta}_0 + \hat{\beta}_T X_{ij} \right) = (y_{ij} - \bar{y}) - (X_{ij} - \bar{X}) \hat{\beta}_T. \quad (51)$$

O S.Q.R. do modelo (48) é dado então por:

$$\hat{u}_r \hat{u}_r = \sum_{i=1}^2 \sum_{j=1}^{n_i} (y_{ij} - \bar{y})^2 - \hat{\beta}_T^2 \sum_{i=1}^2 \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2. \quad (52)$$

Uma vez que (52) coincide com S.Q.R.<sub>Total</sub> do rácio  $F_2$  e (43) com S.Q.R.<sub>dentro</sub> do mesmo rácio, fica demonstrada a igualdade entre esta estatística definida no âmbito da ANCOVA e a que permite conduzir o mesmo teste através de um modelo de regressão com variáveis dummy (estatística (34)).

Se a hipótese  $H_0^3$  for rejeitada, o modelo (38) permite encontrar a equação de regressão para cada grupo de observações através da simples atribuição de valores às variáveis dummy. Mais especificamente,

$$\hat{y}_{ij} = \left( \hat{\delta}_0 + \hat{\delta}_i \right) + \hat{\beta} X \quad (53)$$

$$\hat{y}_{2j} = \hat{\delta}_0 + \hat{\beta} X \quad (54)$$

para os indivíduos que trabalham no horário diurno e nocturno, respectivamente. Substituindo (39) e (40) em (53) e (54) estas últimas transformam-se em

$$y_{ij} = \bar{y}_1 - \bar{X}_1 \hat{\beta} + \hat{\beta} X = \bar{y}_1 - (X - \bar{X}_1) \hat{\beta} \quad (55)$$

$$y_{2j} = \bar{y}_2 - \bar{X}_2 \hat{\beta} + \hat{\beta} X = \bar{y}_2 - (X - \bar{X}_2) \hat{\beta} \quad (56)$$

confirmando assim que um modelo de regressão com variáveis dummy conduz precisamente às mesmas equações de regressão separadas dadas pela fórmula (20) em análise de covariância.

Por fim, se a hipótese  $H_0^3$  não for rejeitada, uma única equação de regressão com todas as observações deverá ser estimada. Essa equação de regressão pode ser

obtida ajustando o modelo (48) que corresponde ao modelo (18) em análise de covariância, como as igualdades (49) e (50) permitem facilmente constatar.

---

## 7. CONSIDERAÇÕES FINAIS

---

Ficou assim demonstrado que a especificação de um modelo de regressão com variáveis quantitativas e variáveis dummy permite conduzir o mesmo tipo de experiências que habitualmente constituem o domínio de aplicação da análise de covariância. Embora utilizando procedimentos distintos, ambas as técnicas conduzem a testes estatísticos equivalentes e a modelos de regressão linear semelhantes. A análise de covariância pode ser estendida a um qualquer número de variáveis quantitativas e variáveis qualitativas. O mesmo se passa obviamente na análise de regressão. Nestes casos, a utilização de software adequado revela-se fundamental.

As vantagens da análise de regressão em relação à análise de covariância são basicamente as mesmas que se verificam numa comparação entre análise de regressão e análise de variância (Valle e Rebelo, 2002): maior facilidade de adaptação a um qualquer número de regressores e/ou de observações e um número acrescido de hipóteses testáveis de uma forma muito simples.

---

## REFERÊNCIAS BIBLIOGRÁFICAS

---

- CHOW, G C. (1960), "Tests of Equality Between Sets of Coefficients in two Linear Regressions", *Econometrica*, Vol. 28(3), pp. 591-605.
- DARLINGTON, R. B. (1990), *Regression and Linear Models*. New York (McGraw-Hill).
- GREENE, W. H. (1993), *Econometric Analysis*. New York (Macmillan Publishing Company).
- GRIFFITHS, W. E., CARTER H. e JUDGE, G. (1992), *Learning and Practicing Econometrics*. New York (John Wiley & Sons, Inc.).
- GUJARATI, D. (1970), "Use of Dummy Variables in Testing for Equality between Sets of Coefficients in Two Linear Regressions: A Note", *The American Statistician*, Vol. 24(1) , pp. 50-52.
- GUJARATI, D. (1970), "Use of Dummy Variables in Testing for Equality between Sets of Coefficients in Two Linear Regressions: A Generalisation", *The American Statistician*, Vol. 24(5) , pp. 18-21.
- HARDY, M. A.(1993), *Regression With Dummy Variables*. Newbury Park (Sage Publications, Inc.)

- HAYS, W. L. (1994), *Statistics*. Philadelphia (Harcourt Brace College Publishers).
- KENNEDY, P. (1986), "Interpreting Dummy Variables", *The Review of Economics and Statistics*, Vol. 68(1), pp. 174-175.
- KOOYMAN, M. A. (1976), *Dummy Variables in Econometrics*. Netherlands (Tilburg University Press).
- KULLBACK, S. e ROSENBLATT H. M. (1957), On the Analysis of Multiple Regression in k Categories, *Biometrika*, 67-83.
- LONG, J. S. e MIETHE, T. D. (1988), *The Statistic Comparison of Groups*. Newbury Park (Sage Publications, Inc.).
- MADDALA, G. S. (1992), *Introduction to Econometrics*. New York (Macmillan Publishing Company).
- MILTON, J. S. e ARNOLD, J. C. (1990), *Introduction to Probability and Statistics, Principles and Applications for Engineering and the Computing Sciences*. New York (McGraw-Hill).
- POKORNI, M. (1992), *An Introduction to Econometrics*. Oxford (Blackwell).
- SANDY, R. (1990), *Statistics for Business and Economics*. New York (McGraw-Hill).
- SEARLE, S. R. e UDELL J. G. (1970), "THE USE OF REGRESSION ON DUMMY VARIABLES IN MANAGEMENT RESEARCH", *MANAGEMENT SCIENCE*, VOL. 16(6), PP. 397-409.
- SILVERTRE, A. (1992), *Análise de variância*. Lisboa (Edições Cedin).
- STEWART, J. (1991), *Econometrics*. Cambridge (Philip Allan).
- SUITS, D. B. (1957), "Use of Dummy Variables in Regression Equations". *Journal of the American Statistical Association*, Vol. 52(280), pp. 548-551.
- SUITS, D. B. (1984), "Dummy Variables: Mechanics V. Interpretation", *The Review of Economics & Statistics*, Vol. 66, pp. 177-180.
- SWAN, C. e CASSIDY H. J. (1975), "One-Time Dummies: Zero-One Dummy Variables Which Take on Only One Non-Zero Value", *Journal of Economics and Business*, Vol. 27(1), pp. 268-271.
- THOMAS, R. L. (1985), *Introductory Econometrics*. London and New York (Longman).
- VALLE, P. O. (1995), *Regressores Dummy e sua Aplicabilidade no Modelo Clássico de Regressão Linear*. Tese de Mestrado.
- VALLE, P. O. E REBELO, E. (2002), "Análise de Variância e Análise de Regressão com variáveis Dummy: Mais Semelhanças do que Diferenças", *Revista de Estatística*, Vol. I, pp. 49-86.
- WONNACOTT, T. J. e WONNACOTT, R. J. (1990). *Introductory Statistics for Business and Economics*. New York (John Wiley & Sons).

## Avaliação de um Módulo de Formação: análise das expectativas e percepções dos formandos

Autores:  
Maria José Conceição  
Andrew B. Hill  
Manuela Magalhães Hill

**VOLUME II**

**2º QUADRIMESTRE DE 2002**

---

## AVALIAÇÃO DE UM MÓDULO DE FORMAÇÃO: ANÁLISE DAS EXPECTATIVAS E PERCEPÇÕES DOS FORMANDOS

---

---

## EVALUATION OF A TRAINING MODULE: ANALYSES OF TRAINEES' EXPECTATIONS AND PERCEPTIONS

---

Autores: Maria José Conceição

- Técnica Superior no Centro de Formação Autárquica de Palmela Luís de Sá.

Andrew B. Hill

- Investigador no DINÂMIA/ISCTE.

Manuela Magalhães Hill

- Professora Associada com Agregação do Instituto Superior de Ciências do Trabalho e da Empresa.

### RESUMO:

- O artigo analisa as expectativas e percepções dos formandos de 14 cursos de qualificação inicial sobre um módulo de formação em Higiene e Segurança no Trabalho. Verificou-se que as expectativas mais comuns eram realísticas e pouco influenciadas pelo género, pela idade e pelas habilitações literárias. A análise de clusters agrupou os curso em quatro grupos de acordo com as expectativas dos formandos. A análise factorial das avaliações de 20 aspectos do módulo mostrou que os formandos perceberam e avaliaram o módulo em termos de seis dimensões - adequação do formador, estrutura do módulo, o seu valor, adequação do ensino, suporte aos formandos e adequação da organização do módulo. Género, idade e habilitações literárias influenciaram significativamente as percepções. Regressão ordinal (PLUM) e regressão convencional mostraram que a avaliação global do módulo obtida através de uma simples escala apenas era influenciada significativamente pelas percepções sobre a estrutura e sobre o valor do módulo. Concluiu-se que análises estatísticas das expectativas e percepções de formandos pode ser um contributo útil na planificação de acções de formação.

### PALAVRAS-CHAVE:

- *Avaliação da formação, higiene e segurança no trabalho, avaliação das expectativas, avaliação das percepções, estatística multivariada.*

### ABSTRACT:

- 191 trainees from 14 courses of initial qualification undertook a training module on health and occupational safety. The article reports statistical analyses of trainee ex-

pectations and perceptions of the module. The most common expectations were found to be realistic and little influenced by gender, age or level of education. However, cluster analysis revealed four groups of initial training courses which differed in terms of trainee expectations. Factor analysis of evaluations of 20 aspects of the module showed that trainees perceived and evaluated it in terms of six dimensions; the adequacy of the formador, the adequacy of teaching, the structure of the module, its value, the support offered to participants and the adequacy of the organization of the module. Perceptions were significantly influenced by gender, age, level of education and type of initial training course. Both ordinal (PLUM) and conventional regression analyses showed that an overall evaluation of the module, made by means of a single rating scale, was significantly influenced only by perceptions of its structure and instrumental value. It is concluded that statistical analyses of trainee expectations and perceptions can provide useful feedback to organisers of training courses.

**KEY-WORDS:**

- *Training evaluation, health and occupational safety, evaluation of expectations, evaluation of perceptions, multivariate statistics.*

## 1. INTRODUÇÃO

É com alguma frequência que os técnicos e gestores de recursos humanos desenharam e implementam acções de formação profissional sem que estas sejam formalmente avaliadas. Contudo, a avaliação de acções de formação assume extrema importância não só na perspectiva de aferir resultados do capital investido, mas sobretudo como contributo para introduzir melhorias.

Actualmente começamos a assistir a uma acumulação de experiências de avaliação e a literatura mais actual mostra uma diversidade de objectos alvo de avaliação. Barbier (1985) faz uma distinção entre os objectos avaliados, em duas vertentes: micro e macro. A primeira pode incidir sobre os agentes, as acções, as medidas de um programa ou projecto de natureza variada. A segunda, reporta-se à avaliação de programas, subprogramas, sistemas, políticas e instituições.

Easterby-Smith (1994) agrupou as principais «escolas de pensamento» sobre a temática da avaliação, sendo de destacar, entre outras, as abordagens de Kirkpatrick (1976), Brinkerhoff (1987) e Meignant (1995).

A metodologia de avaliação proposta por Kirkpatrick tem por base uma tipologia assente em quatro critérios ou níveis de avaliação articulados entre si. O primeiro diz respeito à « reacção » dos formandos, que, de acordo com Harrison (1988) e Newby (1992), é o nível mais focado num processo de avaliação – consiste em auscultar no final da acção de formação, o grau de satisfação dos formandos relativamente a essa acção de formação. O segundo nível, designado por « aprendizagem », refere-se às alterações produzidas pela acção de formação nas atitudes, nos conhecimentos e nas competências dos formandos. O terceiro nível refere-se à « mudança de comportamento em contexto de trabalho » e pressupõe medir a mudança do comportamento no desempenho da função, enquanto que o quarto nível diz respeito ao « impacto organizacional », ou seja, o impacto da formação sobre os objectivos da organização, mais precisamente, na produtividade, nos custos, no absentismo, entre outros.

Meignant (1995) também sugeriu uma perspectiva de avaliação da formação assente em quatro níveis: avaliação da satisfação dos formandos, avaliação dos conhecimentos que os formandos adquiriram através do curso, avaliação do modo como ocorreu a transferência dos conhecimentos para situações de trabalho e avaliação dos efeitos da formação, ou seja, avaliar se esta permitiu alcançar os objectivos inicialmente definidos, quer pelos formandos quer pelas empresas.

Desta breve descrição torna-se claro que os modelos de avaliação apresentados por Kirkpatrick e Meignant são muito semelhantes e que ambos envolvem um tipo subjectivo de avaliação (a satisfação dos formandos). Os outros três tipos de avaliação são relativamente mais objectivos e, do nosso ponto de vista, referem-se a três tipos de utilidade da formação: *utilidade para o desenvolvimento* (desenvolvimento dos(as) conhecimentos/competências do formando), *utilidade prática pessoal* (alterações no desempenho do formando no trabalho) e *utilidade prática organizacional* (alterações no desempenho na organização).

Uma metodologia mais compreensiva de avaliação da formação é a apresentada no modelo de seis estádios sugerido por Brinkerhoff (1987). Embora neste modelo estejam inseridos os três tipos de avaliação objectiva sugeridos por Kirkpatrick e Meignant, o modelo de Brinkerhoff oferece uma aproximação mais detalhada e dinâmica na medida em que integra a avaliação da planificação da acção de formação com a avaliação da sua eficácia. Ou seja, o modelo baseia-se no pressuposto de que a avaliação da formação deve começar por identificar as necessidades de formação que deram origem a objectivos realísticos para a formação (estádio 1). Os objectivos definidos de acordo com as necessidades detectadas são utilizados para conceber o curso e esta concepção deve ser avaliada (estádio 2). O curso, então concebido, deve ser operacionalizado em termos de componentes específicas e programas de formação que também devem ser avaliados (estádio 3). Os resultados imediatos da formação nos formandos são então avaliados em termos de aquisição dos conhecimentos relacionados com os objectivos pedagógicos estabelecidos (estádio 4). O estágio 5 envolve uma avaliação do modo como os conhecimentos e competências adquiridos se transferiram para o desempenho das tarefas e a fase final (estádio 6) avalia o impacto da formação no contexto da organização onde os formandos estão inseridos.

Kirkpatrick (1976) e Meignant (1995) defendem que a avaliação da satisfação dos formandos é um procedimento adequado como primeira fase dum processo de avaliação da formação porque fornece, a quem planeia os cursos, informação útil sobre a adequação do curso em termos dos métodos adoptados, dos conteúdos dos programas, da adequação do formador, da documentação distribuída, etc. No entanto, este ponto de vista pode ser questionado. A «satisfação» dum formando representa o modo como as percepções que o formando tem sobre o curso se identificam com as necessidades ou expectativas que levou com ele para o curso. Estas necessidades/expectativas podem, **ou não**, ser congruentes com os objectivos do curso. Se, tal facto não for relevante, as medidas de satisfação podem fornecer informação enganosa a quem planeia o curso.

Deste modo pensamos que para se conseguir informação útil e pertinente para a concepção de uma acção de formação, é mais vantajoso medir separadamente as expectativas que os formandos trazem para o curso e as percepções que eles têm da formação. Neste contexto, vamos utilizar o termo «percepções» para indicar avaliação estruturada de quão «bom» ou «mau» os formandos classificam determinados aspectos da acção de formação (em vez de quão «satisfeitos» estão ou não com esses aspectos). Conhecer as expectativas dos formandos pode contribuir para o melhor planeamento do curso. Por um lado, permite seleccionar apenas, para dada acção de formação, formandos cujas expectativas sejam congruentes com a natureza e objectivos da acção. Por outro, se a organização do curso tiver conhecimento das expectativas dos formandos, pode delinear uma breve introdução ao curso, que se poderá designar como uma sessão de abertura, em que são transmitidos claramente os objectivos e que clarifique exactamente o que os formandos podem esperar do curso – um aspecto que poderá reduzir a taxa de «abandono». Conhecer o modo como os formandos percebem e avaliam aspectos do curso poderá sugerir modificações necessárias durante a planificação do curso por forma a aumentar a probabilidade de os objectivos do curso se identificarem com a aquisição de conhecimentos e competências. Este tipo de *feedback* deverá trazer melhoria ao que atrás referimos por *utilidade para o desenvolvimento* da acção de formação.

Consideramos que a abordagem à avaliação subjectiva baseada na formação que acima delineámos se aplica, em particular, a acções de formação ministradas por entidades externas e não a acções de formação levadas a cabo pelas próprias empresas, pois estas normalmente baseiam-se nas suas próprias premissas. Isto porque, naquelas, os formandos provêm de empresas diferentes, de diferentes sectores de actividade podendo representar uma variedade de profissões diferentes (como é o caso da situação apresentada no presente estudo). Em tais circunstâncias, é natural que a diversidade de experiências de trabalho esteja associada a diferentes expectativas sobre o curso e a diferentes percepções sobre o mesmo. Em Portugal não é invulgar encontrarmos acções ou módulos de formação ministrados por entidades formadoras externas com formandos muito heterogéneos. Por exemplo, num estudo recente efectuado sobre uma amostra representativa constituída por 687 empresas da indústria transformadora localizadas nas áreas metropolitanas de Lisboa e Porto, 41,1% das empresas ofereceram acções de formação aos seus trabalhadores durante o período 1996-1998 sendo que 72,1% destas utilizaram entidades externas para darem a formação (Almeida, Fernandes e Hill, 2000).

É razoável esperar que em acções de formação onde os formandos diferem em termos de profissão e actividade onde trabalham, estes também difiram num conjunto de outras características, tais como, género, idade ou nível de habilitação. Quando isto acontece, podemos pensar, por exemplo, que diferenças individuais desta natureza poderão influenciar as expectativas e as percepções dos formandos sobre a acção de formação. É óbvio que tais influências devem ser tidas em conta numa avaliação da formação porque podem fornecer informação útil para os organizadores do curso. Por exemplo, as percepções feitas pelos formandos mais velhos poderão sugerir que a estrutura do curso deve incluir maior suporte para este tipo de formando.

Este artigo apresenta os resultados de um estudo sobre expectativas e percepções dos formandos associadas com um **Módulo** de formação em Saúde, Ambiente e Segurança no Trabalho, levado a cabo num Centro de Formação Profissional do I.E.F.P., e comum a todos os cursos de qualificação inicial. Deste modo, estamos perante um conjunto de formandos muito heterogéneos em termos de formações profissionais.

Os objectivos específicos do estudo eram os seguintes:

### **Objectivo 1**

Conhecer a natureza das expectativas mais comuns que os formandos tinham sobre o Módulo.

#### **Objectivo 1(a)**

Determinar se o género, a idade, o nível de habilitação e o tipo de curso que frequentavam influenciaram as expectativas dos formandos sobre o Módulo.

### **Objectivo 2**

Determinar qual a percepção que os formandos tinham sobre o Módulo. Isto é, determinar as dimensões em termos das quais os formandos avaliaram o Módulo.

### **Objectivo 2(a)**

Determinar a influência que o género, a idade e o nível de habilitação tiveram sobre as percepções dos formandos sobre o Módulo.

### **Objectivo 2(b)**

Determinar a influência do curso de formação inicial sobre as percepções dos formandos sobre o Módulo.

### **Objectivo 3**

Determinar se as percepções dos formandos estavam relacionadas com a avaliação global que fizeram sobre o Módulo.

---

## **2. MÉTODO**

### *A natureza e duração do Módulo*

O módulo de formação em Saúde, Ambiente e Segurança no Trabalho está inserido em todos os 14 cursos de qualificação inicial e integrado na componente sócio cultural. A duração é de 20 horas, distribuídas por cinco sessões de três horas e meia e uma sessão de duas horas e meia, duas vezes por semana, em horário laboral - durante as manhãs. Contou com a participação de um formador, que nesta área é sempre externo, e com um número previsível de formandos que oscila entre 12 e 14 formandos por acção. Este módulo está previsto ser ministrado numa sala do centro, dispondo à partida de alguns equipamentos e materiais tais como retroprojector, vídeo e televisor, quadro de papel e de porcelana.

### *Os dados*

Os dados foram recolhidos por meio de questionário aplicado a todos os formandos que frequentaram os cursos de qualificação inicial em 1998. Para além de perguntas de caracterização geral, o questionário incluiu, entre outras, perguntas sobre motivos de inscrição no curso, sobre verificação das expectativas pessoais, sobre objectivos do módulo e sobre apreciação/avaliação do módulo

### *Os formandos*

O número de formandos que frequentaram em 1998 o Módulo de Saúde, Ambiente e Segurança no Trabalho totalizou 191. O questionário foi aplicado no final do módulo a todos os formandos. A maioria dos formandos (84,7%) inscreveu-se no módulo por iniciativa própria, 5,8% por sugestão do chefe e 9,5% por sugestão de outra pessoa (não especificada). Havia um predomínio de formandos do sexo feminino (64,4%). A estrutura etária era bastante jovem, sendo que 63,6% tinham menos de 30 anos (só uma pessoa tinha mais de 54 anos). A maioria dos formandos (59,7%) tinha escolaridade até ao 9º ano.

Os 191 formandos estavam a frequentar um dos 14 cursos de qualificação inicial, sendo os cursos de Esteticismo e de Ajudantes de Cabeleireiro os que detinham

maior número de formandos. O Quadro 1 mostra a natureza dos cursos bem como a repartição dos formandos por género. Note-se que 3 formandos não referiram o género.

#### *A análise dos dados*

Os dados foram analisados por meio das seguintes análises estatísticas: teste do Qui-quadrado, análise de clusters, ANOVA, MANOVA, análise factorial e análise de regressão. Os detalhes sobre as técnicas encontram-se na secção seguinte.

### 3. RESULTADOS

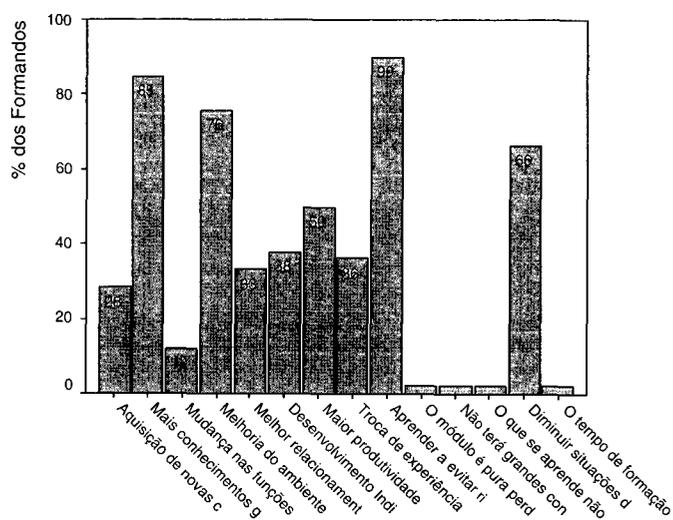
#### **Objectivo 1. Expectativas dos formandos sobre o módulo**

No Gráfico 1 podemos visualizar as expectativas específicas sobre o módulo que os formandos consideraram mais e menos importantes de entre 14 aspectos apresentados. O que os alunos mais esperavam do módulo era «aprender a evitar riscos», expectativa referida por 90% dos formandos. Seguidamente surgem «mais conhecimentos gerais» (85%), «melhoria do ambiente de trabalho» (76%), «diminuir situações desagradáveis» (66%) e, com apreciável menos importância «maior produtividade» (50%). É interessante notar que apenas 28% dos formandos pensavam vir a «adquirir novas competências» e apenas pouco mais de um terço (38%) esperava conseguir «desenvolvimento individual». Contudo, só menos de 5% dos formandos pensava que o módulo era «pura perda de tempo» e que «não terá grandes consequências».

**Quadro 1**  
**Repartição dos formandos por curso e género**

		Género		Total
		Masculino	Feminino	
Curso	Esteticismo		36	36
	Op. Máq. Ferramentas	12		12
	Soldadura	9		9
	Práticas Administrativas	1	10	11
	Jardinagem	5	5	11
	Costura Industrial		16	16
	Cabeleireiro	1	25	26
	Manicura/Pedicura		8	8
	Contabilidade e Gestão	4	5	10
	Secretariado Comercial		12	12
	Serralharia Civil	8		8
	metalomecânica II	11		11
	Práticas técnico-comerciais	7	4	12
	Serralharia mecânica	9		9
Total		67	121	191

**Gráfico 1**  
**Expectativas dos formandos**



### **Objectivo 1(a). A influência do género, da idade do nível de habilitação e do curso nas expectativas dos formandos**

#### *A influência do género sobre as expectativas*

A aplicação de testes de independência do qui-quadrado revelaram que, exceptuando a «aquisição de mais conhecimentos gerais», o género do formando não tinha influência estatisticamente significativa sobre as expectativas. Neste caso, as expectativas eram mais frequentes ( $P = 0,044$ ) nas mulheres (88,4%) que nos homens (77,3%).

#### *A influência da idade e das habilitações literárias nas expectativas dos formandos*

A aplicação do teste de independência do qui-quadrado mostrou que a idade só influenciava significativamente a expectativa «troca de experiências» ( $P = 0,04$ ). Com efeito, ela foi referida por 31,6% dos formandos mais novos (<30 anos) em comparação com 47,0% dos mais velhos ( $\geq 30$  anos).

O nível de habilitação influenciava significativamente apenas a expectativa «mudança nas funções desempenhadas». Esta expectativa observou-se em 16,3% dos formandos com habilitações até ao 9º ano, em comparação com 6,5% dos que fizeram o 10º ano ou mais.

#### *A influência do curso nas expectativas dos formandos*

O número de casos da amostra foi insuficiente para se poder determinar se as expectativas que os formandos (191) tinham antes de frequentar o módulo eram significativamente diferentes entre cursos (14). No entanto, pensando que dentro do total de formandos deverão existir subgrupos de cursos significativamente diferentes no que diz respeito às expectativas trazidas pelos formandos, foi efectuada uma análise de clusters considerando os 14 cursos como casos e as expectativas como variáveis.

A análise suportou esta hipótese pois distinguiu 4 clusters ou grupos de cursos. A natureza destes grupos encontra-se no Quadro 2.

**Quadro 2**  
**Resultado da análise de clusters**

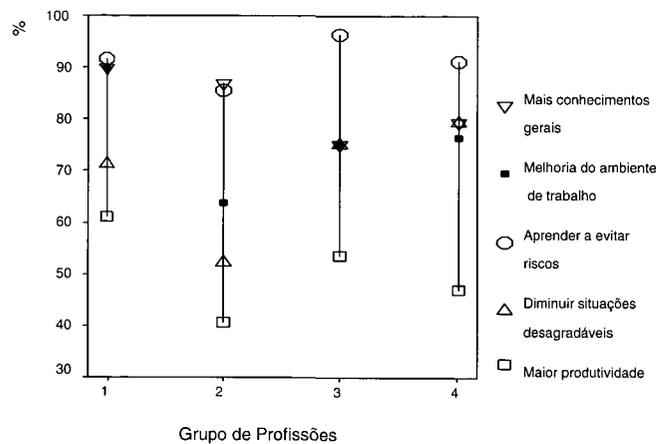
Grupo	Cursos dentro do Grupo	Número de Formandos no Grupo	% Homens	% Idade < 30	% Hab. Lit. <10ºano
1	Esteticismo Práticas administrativas Secretariado comercial	59	1,7	81,4	1,8
2	Op. Maquinas ferramentas Soldadura Jardinagem Costura industrial Contabilidade e Gestão Práticas técnico-comerciais	70	55,2	42,2	76,9
3	Serralharia civil Serralharia mecânica Metalomecânica II	28	100	60,7	96,3
4	Cabeleireiro Manicura/Pedicura	34	2,9%	75,8%	84,8%

Os grupos 1 e 4 representam cursos «tradicionalmente femininos» e o grupo 3 representa cursos «tradicionalmente masculinos». O grupo 4 é um grupo misto de cursos de «colarinho branco e colarinho azul» no qual ambos os géneros se encontram representados.

Como se pode observar no gráfico 2, os grupos diferem em termos das cinco expectativas mais vezes referidas. No entanto, três destas cinco expectativas não diferiram significativamente entre os grupos. Estas foram «aquisição de mais conhecimentos gerais» ( $P = 0,236$ ), «aprender a evitar riscos» ( $P = 0,386$ ) e «maior produtividade» ( $P = 0,134$ ). As diferenças significativas verificaram-se nas expectativas «melhorar o ambiente» ( $P = 0,008$ ) e «diminuir situações desagradáveis» ( $P = 0,016$ ).

Gráfico 2

Expectativas mais vulgares por grupo de profissões

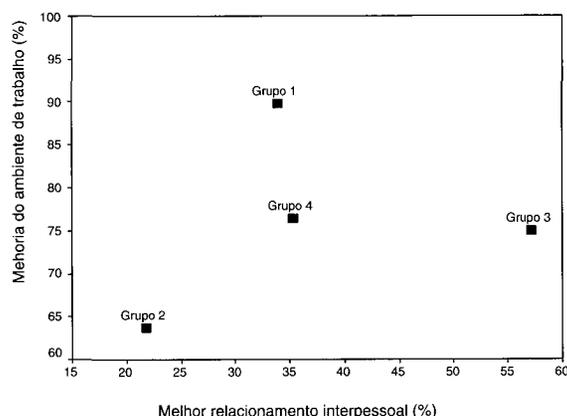


Das expectativas referidas por menos de 50% dos formandos, apenas duas diferiram significativamente entre os quatro grupos de cursos. Estas eram: «aquisição de novas competências» ( $P = 0,015$ ) e «melhor relacionamento interpessoal» ( $P = 0,010$ ).

Mas, de todas estas expectativas, o que mais claramente distinguiu cada um destes grupos de cursos de cada um dos outros foi a expectativa «melhoria do ambiente de trabalho» e «melhor relacionamento interpessoal». O gráfico 3 mostra as diferenças entre os quatro grupos de cursos em termos da percentagem de formandos que manifestaram estas duas expectativas.

Para além destas diferenças significativas, os grupos também registam diferenças apreciáveis em termos de género, idade e escolaridade, como mostra o Quadro 2. No entanto, as análises anteriormente efectuadas sobre a influência destas características nas expectativas dos formandos mostraram que nenhuma delas influenciava significativamente nem a «melhoria do ambiente de trabalho» nem o «melhor relacionamento interpessoal». Deste modo, é possível concluir que as diferenças que se verificaram entre os quatro grupos em termos destas duas expectativas não dependiam por si só de nenhuma das três características dos formandos consideradas. No entanto, tendo em conta as diferenças em género, idade e escolaridade dos formandos, é possível descrever os quatro grupos da seguinte forma:

**Gráfico 3**  
**Percentagem de formandos que têm as expectativas «Melhoria do ambiente de trabalho» e «Melhor relacionamento interpessoal»**  
(analisada por grupo)



**Grupo 1 – Cursos tradicionalmente femininos mais exigentes** (mais exigentes porque quase todos os formandos têm pelo menos o 10º ano). É o grupo mais expectante relativamente à «melhoria do ambiente de trabalho» e com expectativa média em relação a «melhor relacionamento interpessoal», quando comparado com os outros três grupos.

**Grupo 2 – Mistura de cursos «colarinho branco/colarinho azul»**, contendo tanto formandos de ambos os géneros. Este grupo é o que tem menos expectativas em relação à «melhoria do ambiente de trabalho» e ao «melhor relacionamento interpessoal», quando comparado com os outros três grupos.

**Grupo 3 – Cursos tradicionalmente masculinos menos exigentes** (menos exigentes porque quase todos os formandos têm menos que o 10º ano). Este grupo tem uma expectativa média em relação à «melhoria do ambiente de trabalho» e a maior expectativa relativamente a conseguir «melhor relacionamento interpessoal», quando comparado com os outros três grupos.

**Grupo 4 – Cursos tradicionalmente femininos menos exigentes** (menos exigentes porque quase todos os formandos têm menos que o 10º ano). Este grupo tem uma expectativa média tanto em relação à «melhoria do ambiente de trabalho» como a «melhor relacionamento interpessoal», quando comparado com os outros três grupos.

## **Objectivo 2. Percepções dos formandos sobre o módulo**

Para analisar as percepções dos formandos sobre o módulo foi aplicada uma análise factorial a 20 itens de uma questão respondida pelos formandos no final do módulo. Estes itens diziam respeito a uma variedade de aspectos do módulo (ver Quadro 3) e foram codificados numa escala de 1 a 5 que vai do «mau» ao «muito bom».

Inicialmente usou-se os 22 itens incluídos no questionário mas dois deles foram removidos da análise factorial porque apresentavam valores baixos em termos de adequabilidade da amostragem na matriz de correlação anti-imagem. O método de extracção seleccionado foi o de máxima verosimilhança e a medida de KMO indica que a amostragem é adequada (0,894). A análise produziu 6 factores que explicaram 70,81% da variância total e o ajustamento aos dados foi adequado ( $\chi^2 = 103,15$ ;  $df = 85$ ,  $P = 0,088$ ). Apenas 5% das correlações residuais eram maiores que 0,05 e nenhuma destas excedeu 0,1. Duas soluções foram obtidas, uma usando rotação ortogonal (varimax) e outra usando rotação oblíqua (oblimin). A experimentação mostrou que uma solução oblíqua com  $\delta = 0,2$  foi a que forneceu melhor aproximação a uma estrutura simples, no entanto, esta solução resultou pouco melhor que a solução ortogonal. Perante isto, a solução ortogonal foi retida por oferecer uma interpretação mais clara dos factores. Os *scores* dos factores para cada formando foram calculados usando o método dos mínimos quadrados. O Quadro 3 apresenta os pesos (*loadings*) das variáveis (ou seja, dos 20 itens do questionário) nos seis factores. Para clarificar a interpretação, foram omitidos os pesos com valor inferior a 0,3.

**Quadro 3**  
**Solução da análise factorial após rotação ortogonal**

VARIÁVEL	FACTOR					
	1	2	3	4	5	6
1. Interesse dos temas	,309		,529	,343		
2. Utilidade dos temas		,429	,455	,317		
3. Aplicabilidade dos temas na sua função		,614	,339			
4. Articulação teoria/prática		,654				
5. Adequação dos exercícios práticos realizados		,686				
6. Clareza de exposição do monitor	,499			,638		
7. Clareza das matérias tratadas	,305			,712		
8. Domínio dos temas pelo monitor	,530			,497		
9. Cumprimento do programa						,660
10. Linguagem utilizada	,546					,371
11. Capacidade do monitor para motivar participantes	,882					
12. Relação com os participantes	,693					
13. Disponibilidade do monitor para clarificar dúvidas	,515				,483	
14. Métodos e meios audiovisuais utilizados		,468			,478	
15. Qualidade da documentação distribuída					,593	
16. Apoio administrativo					,415	
17. Duração do módulo		,415				
18. Aquisição novos conhecimentos/métodos/técnicas			,780			
19. Coordenação do curso			,495			
20. Nível de participação permitido	,357				,335	,345

Aos factores foi atribuída a seguinte interpretação:

- Factor 1. **Adequação do monitor/ formador**
- Factor 2. **Estrutura do módulo**
- Factor 3. **O valor do módulo**
- Factor 4. **Adequação do ensino**
- Factor 5. **Suporte aos formandos**
- Factor 6. **Adequação da organização do módulo**

Não estamos muito confiantes com a interpretação dada ao factor 6 porque embora o item 9 seja o que apresenta peso mais elevado, a inclusão dos itens 10 e 20 na formação do factor sugere que este mede mais do que simplesmente «cumprimento do programa». O factor parece representar a adequação do modo como o módulo foi organizado para dar cumprimento ao programa (isto, através do tipo de linguagem utilizada e do nível de participação permitido).

Tendo em atenção que todos os pesos nos factores eram positivos e que pontuações numéricas mais elevadas dos 20 itens no questionário representavam avaliações positivas, valores elevados em cada factor implicam percepções mais positivas (i.é, mais favoráveis) enquanto que valores mais baixos implicam percepções menos positivas (menos favoráveis).

#### **Objectivo 2(a). A influência do género, da idade e do nível de habilitação nas percepções sobre o módulo.**

Para analisar a influência das três variáveis foi utilizada uma MANOVA, tendo os scores dos seis factores ortogonais como variáveis dependentes e o género, a idade (< 30 / ≥ 30) e as habilitações literárias (< 10º ano / ≥ 10º ano) como variáveis independentes. Os valores de F dos testes multivariados resultaram significativos para os três efeitos principais, género (P = 0,047, potência = 0,761), idade (P = 0,012, potência = 0,876) e habilitações literárias (P = 0,011, potência = 0,885). Só uma das interacções de primeira ordem (género × nível de habilitações) (P = 0,009, potência = 0,892) resultou significativa e não se verificou interacção de segunda ordem.

##### *O efeito do género*

Os efeitos entre os grupos mostram que apenas se observaram diferenças significativas entre os homens e as mulheres para as variáveis dependentes «estrutura do módulo» (P = 0,029, potência = 0,591) e «adequação da organização do módulo» (P = 0,008, potência = 0,764). Em média, as mulheres perceberam a estrutura do módulo de maneira mais positiva que os homens. Esta afirmação deve, no entanto, ser lida com cuidado porque a potência do teste era apenas modesta e a variável género estava envolvida num termo de interacção significativo. Em média, as mulheres também perceberam a adequação da organização do módulo mais favoravelmente que os homens.

### *O efeito da idade*

Em média, os formandos mais novos tinham uma percepção significativamente mais favorável que os mais velhos (>30 anos) sobre a «estrutura do módulo» ( $P = 0,002$ , potência = 0,871) e sobre o «suporte aos formandos» ( $P = 0,018$ , potência = 0,663).

### *O efeito das habilitações literárias*

Os formandos com menos que o 10º ano de escolaridade apresentaram em média percepções mais positivas sobre a «estrutura do módulo» que os que têm o 10º ano ou mais ( $P < 0,0001$ , potência = 0,948). Muito embora se tenha observado uma potência elevada e um bom nível de significância, temos de ter em atenção que existe uma interacção entre esta variável e variável género, a qual deve ser explorada.

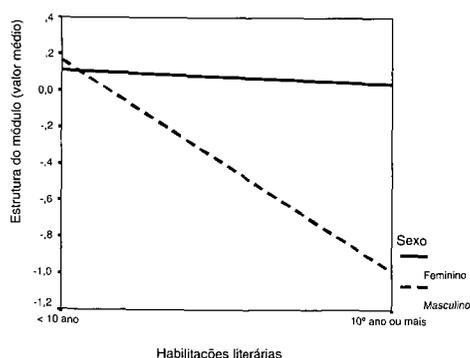
### *O efeito conjunto do género e habilitações literárias*

A interacção entre género e habilitações literárias apresentou-se significativa para a «estrutura do módulo» ( $P = 0,019$ , potência = 0,653) e quase significativa para a «adequação do ensino» ( $P = 0,051$ ) embora a potência não impressione (potência = 0,498). A natureza destas duas interacções pode ser mais facilmente entendida através dos gráficos 4 e 5.

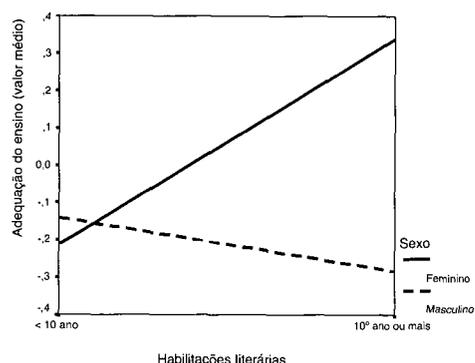
O gráfico 4 mostra que, no grupo dos formandos com habilitações literárias até ao 9º ano, não se registou grande diferença entre homens e mulheres quanto às percepções médias sobre a «estrutura do módulo» - ambos apresentam percepções ligeiramente positivas. Mas, no grupo dos formandos com habilitações literárias superiores, a percepção dos homens sobre a estrutura do módulo foi, em média, significativamente menos favorável que a das mulheres.

No que diz respeito à «adequação do ensino», o gráfico 5 mostra uma situação semelhante. Também no grupo dos formandos com habilitações literárias até ao 9º ano, não há grande diferença entre homens e mulheres quanto às percepções médias sobre o módulo - ambos apresentam percepções ligeiramente negativas. E, no grupo com habilitações literárias superiores, a percepção dos homens sobre a estrutura do módulo foi, em média, significativamente menos favorável que a das mulheres.

**Gráfico (4)**  
**Interacção género x habilitação**  
**na estrutura do módulo**



**Gráfico (5)**  
**Interacção género x habilitação**  
**na adequação do ensino**



### Objectivo 2(b): A influência do curso nas percepções sobre o módulo

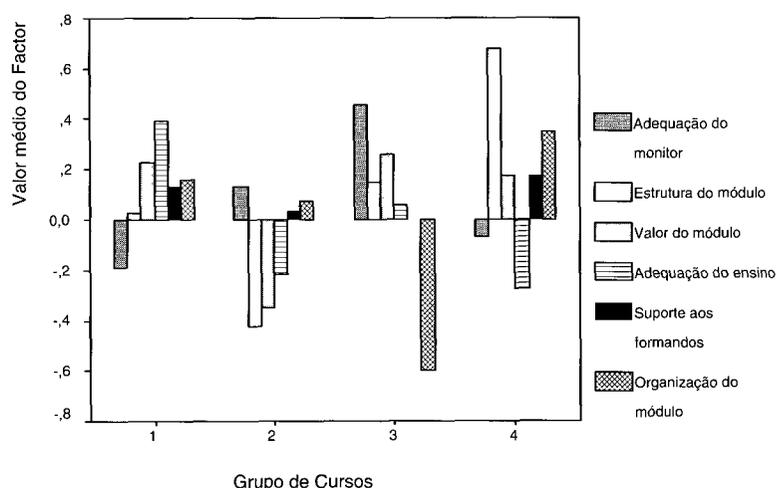
Como o número de casos é insuficiente para se poder comparar os 14 cursos de iniciação profissional, utilizou-se a MANOVA para comparar os quatro clusters, ou seja, os quatro grupos de cursos (ver Quadro 2) em termos dos factores ortogonais.

O valor de  $F$  para o teste multivariado foi significativo ( $P < 0,0001$ ) e a análise dos efeitos entre grupos mostrou diferenças significativas entre os grupos em cinco dos seis factores. Estes cinco factores são percepções sobre a «adequação do monitor» ( $P = 0,05$ , potência = 0,640), sobre a «estrutura do módulo» ( $P < 0,0001$ , potência = 0,999), sobre o «valor do módulo» ( $P = 0,002$ , potência = 0,930), sobre a «adequação do ensino» ( $P = 0,001$ , potência = 0,940) e sobre a «adequação da organização do módulo» ( $P < 0,0001$ , potência = 0,983).

O gráfico 6 apresenta os valores (*scores*) médios dos factores em cada um dos quatro grupos de cursos. Em média, os grupos 2 (misto de cursos «colarinho branco/colarinho azul») e 3 (cursos tradicionalmente masculinos menos exigentes) fazem percepções mais positivas que os grupos 1 (cursos tradicionalmente femininos mais exigentes) e 4 (cursos tradicionalmente femininos menos exigentes) sobre a «adequação do monitor». Contudo, os testes *post hoc* mostraram que a diferença só era significativa ( $P = 0,001$ ) entre os grupos 1 e 3.

Os formandos dos grupos 1, 3 e 4 (cursos tradicionalmente femininos ou tradicionalmente masculinos) perceberam mais favoravelmente a «estrutura do módulo» e o «valor do módulo» que os do grupo 2, e os testes *post hoc* mostraram que, em ambos os casos, o grupo 2 diferia significativamente dos outros três grupos ( $P < 0,05$  ou melhor).

**Gráfico 6**  
Valores médios das percepções (factores) por grupo de cursos



Em média, os formandos dos grupos 1 e 3 perceberam a «adequação do ensino» mais positivamente que os dos grupos 2 e 4, mas os testes *post hoc* mostraram que apenas o grupo 1 diferia significativamente dos grupos 2 e 4 ( $P < 0,05$  ou melhor).

Em média, os quatro grupos não diferiam significativamente nas suas percepções em relação ao «suporte aos formandos», todos os grupos apresentaram percepções médias neutras ou ligeiramente positivas.

Sobre a «adequação da organização», os grupos 1, 2 e 4 tinham percepções médias mais positivas que o grupo 3 mas as diferenças apenas eram estatisticamente significativas entre os grupos 3 e 1 e entre os grupos 3 e 4 ( $P < 0,025$  ou melhor).

### **Objectivo 3: Relação entre percepções dos formandos e a avaliação global sobre o módulo**

A avaliação do módulo de formação na sua globalidade foi obtida a partir de uma única questão colocada aos formandos, usando uma escala com quatro níveis. Tem-se que 18,3% dos formandos avaliaram o módulo como «muito bom», 34,0% como «bom», 22,0% como «médio» e 2,1% como «insuficiente». Infelizmente, 23,6% dos formandos não responderam a esta questão. Analisando as características destes 23,6% de formandos verifica-se que 60,5% eram do sexo masculino, 53,5% tinham 30 ou mais anos, 76,7% tinham menos que o 10º ano de escolaridade e 64,4% estavam a frequentar cursos de iniciação profissional do grupo 2. Como esta percentagem de não respostas é bastante elevada, os resultados a seguir apresentados devem ser aceites com uma certa reserva. Contudo, mais adiante iremos tecer mais comentários sobre estas não respostas.

Como a variável «avaliação global do módulo» está medida numa escala ordinal, efectuou-se uma regressão ordinal (PLUM) usando como variável dependente a avaliação global do módulo e, como variáveis independentes, os seis factores ortogonais das percepções sobre o módulo. Os resultados mostraram que apenas duas das seis percepções estavam significativamente relacionadas com a avaliação global – a percepção sobre a «estrutura do módulo» ( $P < 0,0001$ ) e a percepção sobre o «valor do módulo» ( $P = 0,01$ ). No entanto, a regressão baseou-se numa matriz de dados com 74,5% de células vazias, questão que levanta sérias dúvidas sobre a validade do modelo. Para testar estes resultados, aplicou-se uma regressão linear clássica aos mesmos dados. Esta mostrou uma correlação múltipla significativa com o valor de 0,430 ( $P < 0,0001$ ) entre a avaliação global e os seis factores, sendo que só as duas percepções «estrutura do módulo» ( $\beta = 0,307$ ,  $P = 0,001$ ) e «valor do módulo» ( $\beta = 0,219$ ,  $P = 0,016$ ) apresentaram coeficientes de regressão parcial significativos. As correlações *part* para estas duas variáveis eram 0,303 e 0,214 e não surgiu evidência de multicolinearidade pois as variáveis independentes eram os factores que resultaram de uma análise factorial com rotação ortogonal. Os gráficos *P-P* dos resíduos estandardizados eram adequadamente lineares e, exceptuando o facto de a variável dependente ser ordinal em vez de métrica, é possível afirmar que estamos perante uma regressão convencional aceitável. Os resultados desta regressão confirmaram os obtidos através da regressão ordinal e indicaram que, neste caso, as percepções sobre a estrutura tinham uma contribuição ligeiramente mais forte na avaliação global que as percepções sobre o valor do módulo. Com efeito, se se efectuar uma regressão linear usando apenas os dois factores «estrutura do módulo» e «valor do módulo» como variáveis independentes, o coeficiente de correlação múltipla passa a ser 0,404 ( $P < 0,0001$ ).

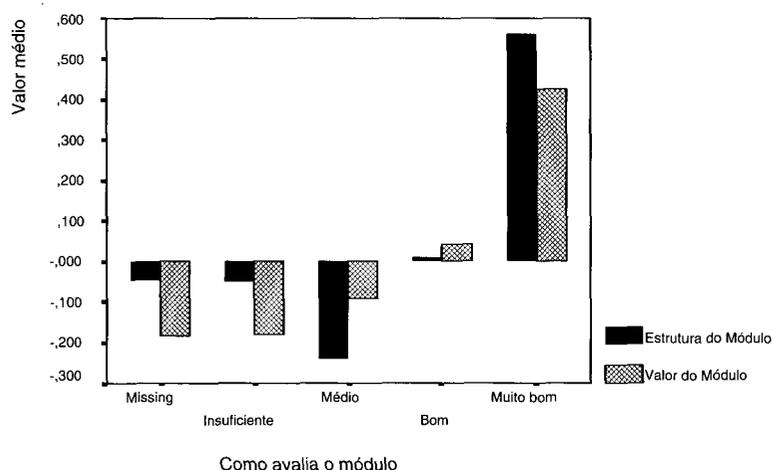
O gráfico 7 mostra, em mais detalhe, a relação entre a avaliação global do módulo e as percepções sobre a estrutura e sobre o valor do módulo.

Os formandos que não avaliaram o módulo na globalidade (grupo «Missing» no gráfico 7), apresentam percepções médias sobre a estrutura (-0,045) e sobre o valor (-0,182) quase idênticas às dos formandos que avaliaram o módulo como tendo sido «insuficiente» (-0,049 para a estrutura e -0,181 para o valor).

Em termos de características pessoais, os formandos que não responderam à questão sobre a avaliação global do módulo são também semelhantes aos quatro formandos que responderam «insuficiente» ao avaliarem o módulo na sua globalidade. Neste último grupo, 3 dos 4 formandos eram homens com menos de 30 anos, com escolaridade inferior ao 10º ano e frequentavam cursos do grupo 2. Esta evidência sugere que é muito provável que os formandos que não fizeram uma avaliação global do módulo, consideravam-no insuficiente mas não tiveram coragem de o dizer.

O gráfico 7 mostra que, em geral, as avaliações globais do módulo mais favoráveis estavam associadas a percepções mais favoráveis sobre a estrutura e sobre o valor do módulo. Encontrou-se, contudo, uma exceção a esta tendência, os formandos que avaliaram o módulo como «insuficiente» foram os que tiveram a percepção menos favorável do valor do módulo (como esperado), no entanto, a sua percepção sobre a estrutura foi mais favorável que a dos formandos que atribuíram uma classificação «média».

**Gráfico 7**  
**Avaliação global e scores médios da estrutura e valor do módulo**



#### 4. DISCUSSÃO

##### *As expectativas dos formandos sobre o Módulo*

O que os formandos mais esperavam do módulo (referido por cerca de dois terços ou mais dos formandos) era que este lhes trouxesse mais conhecimentos gerais e vários tipos de conhecimentos específicos relacionados com o evitar riscos, a melhoria do ambiente de trabalho e a diminuição de situações desagradáveis.

Só uma minoria de formandos esperavam que o módulo tivesse especificamente efeitos pessoais no desenvolvimento individual (30%), na mudança nas funções desempenhadas (12%) e no relacionamento interpessoal (33%). Deste modo, para um módulo de Saúde, Ambiente e Segurança no Trabalho, as expectativas mais vezes referidas parecem ser *realísticas* e congruentes com os objectivos do curso. Observe-se, no entanto, que só pouco mais de um quarto dos formandos (28%) é que esperava que o módulo lhes desse novas *competências*. Se relacionarmos este facto com as expectativas mais referidas, que estão relacionadas com a aquisição de *conhecimentos*, parece que os formandos, ao formularem as suas expectativas, fizeram uma clara distinção entre aquisição de conhecimentos e aplicação desses conhecimentos (porque «competência» é essencialmente a aplicação prática e adequada do conhecimento). Isto sugere que talvez os formandos não esperassem que o conteúdo do módulo fosse suficientemente prático que lhes permitisse virem a aplicar os conhecimentos práticos adquiridos. Mas, se notarmos que as percepções mais positivas sobre a estrutura do módulo envolveram avaliações mais positivas sobre a adequação dos exercícios realizados, a articulação teoria/prática e a aplicabilidade dos temas na sua função, e que as percepções mais positivas sobre o valor do módulo envolveram avaliações mais positivas sobre a utilidade dos temas e a aplicabilidade dos temas na sua função, podemos talvez inferir que, muito embora os formandos esperassem que o módulo lhes fornecesse novos conhecimentos em vez de novas competências, é provável que aqueles que consideraram de maneira mais positiva a estrutura e o valor do módulo tenham efectivamente adquirido novas competências.

#### *Os factores pessoais que influenciam as expectativas do Módulo*

Como esperado, encontrámos diferenças consideráveis entre os formandos no que se refere ao género, à idade e às habilitações literárias que possuíam. Contudo, estes factores praticamente não influenciaram as expectativas dos formandos sobre o módulo. O único resultado estatisticamente significativo encontrado foi que as mulheres esperavam adquirir mais conhecimentos, muito embora esta expectativa tenha sido referida por mais de três quartos dos homens e das mulheres. Nem a idade nem o nível de habilitação mostrou influência significativa sobre as expectativas mais comuns. Embora à primeira vista estes resultados possam parecer surpreendentes, eles podem provavelmente ser explicados pelo conteúdo específico do módulo. Um módulo de formação em Saúde, Ambiente e Segurança no Trabalho aborda temas muito importantes para o bem estar do trabalhador e por isso é de interesse comum a todos os (futuros) trabalhadores, ou seja, aos formandos de todos os cursos. Deste modo, é natural que as expectativas mais comuns (que são realistas) sobre um módulo deste tipo sejam formuladas por formandos de ambos os sexos, de todas as idades e níveis de habilitação.

Quatro grupos de cursos diferiram entre si em termos de uma das expectativas mais referidas pelos formandos (melhoria do ambiente de trabalho) e uma das expectativas menos referidas pelos formandos (melhor relacionamento interpessoal). Os formandos dos cursos tradicionalmente femininos mais exigentes (esteticismo, práticas administrativas e secretariado comercial) tinham maiores expectativas sobre a melhoria do ambiente de trabalho, embora a expectativa fosse apenas média em relação a melhorar o relacionamento interpessoal. Os formandos do grupo que mistura cursos de colarinho branco com cursos de colarinho azul tinham as menores expectativas de todas relativamente aos dois aspectos.

Os formandos dos cursos tradicionalmente masculinos menos exigentes (Serralharia/metalmecânica) tinham expectativas médias sobre a melhoria do ambiente do trabalho e a maior expectativa sobre melhor relacionamento interpessoal, e os formandos dos cursos tradicionalmente femininos menos exigentes (Cabeleireiro/Manicura-Pedicura) tinham expectativas moderadas tanto sobre a melhoria do ambiente de trabalho como em relação a melhor relacionamento interpessoal.

Embora os formandos nos quatro grupos de cursos difiram apreciavelmente em termos de género, idade e escolaridade, as diferenças de expectativas entre os grupos não pode ser explicada como resultado **directo** de diferenças nestas características. Isto porque nenhuma das expectativas que diferia entre os grupos era influenciada significativamente nem pela idade, nem pelo género nem pelo nível de habilitação. Mais precisamente, género, idade e nível de habilitação não influenciavam directamente nem a «melhoria do ambiente de trabalho» nem o « melhor relacionamento interpessoal» mas, quando combinados com o tipo de curso, poderão ter alguma influência **indirecta**. Infelizmente, os dados disponíveis não permitiram explorar esta hipótese.

#### *As percepções dos formandos sobre o Módulo*

A análise factorial mostrou que as relações entre os 20 aspectos do módulo avaliados tinham subjacentes seis percepções (dimensões de avaliação) estruturadas sobre o módulo. Este resultado é interessante pois veio salientar algumas **distinções** feitas pelos formandos, distinções essas difíceis de captar a partir dos 20 aspectos iniciais. Por exemplo, os formandos tiveram percepções diferentes sobre a **adequação do monitor/formador** e a **adequação do ensino**. A primeira diz fundamentalmente respeito à disponibilidade do monitor, domínio dos temas, linguagem utilizada, capacidade para motivar participantes, etc., e a segunda, à clareza de exposição, clareza das matérias tratadas e interesse e utilidade dos tópicos ensinados. Analogamente, os formandos fizeram distinção entre **estrutura do módulo** e **organização do módulo**. A estrutura foi percebida em termos da aplicabilidade dos temas na função, do equilíbrio teoria/prática, da duração do módulo e dos métodos de ensino utilizados. A organização foi principalmente percebida em termos do cumprimento do programa mas também em termos da linguagem utilizada e do nível de participação permitido. O módulo também foi percebido em termos de **suporte aos formandos**, o qual engloba o apoio administrativo, a qualidade de documentação distribuída, os métodos e meios audiovisuais utilizados e a disponibilidade do monitor para clarificar dúvidas. Mas talvez o aspecto mais interessante que emerge da análise factorial foi verificar que os formandos parecem avaliar o **valor do módulo** principalmente em termos do seu valor instrumental. Ou seja, em termos da sua utilidade, aplicabilidade para a função e aquisição de novos conhecimentos, métodos e técnicas (que apenas 28% dos formandos esperavam que o módulo lhes fornecesse).

#### *Influência do género, da idade e das habilitações literárias sobre percepções do Módulo*

Tanto o género como a idade ou as habilitações literárias têm pouca influência sobre as percepções que os formandos têm do módulo. Os mais novos valorizaram mais positivamente que os mais velhos tanto a estrutura do módulo como o suporte aos formandos mas nenhuma das outras percepções foi influenciada pela idade. Estes

resultados podem fornecer informação útil para os gestores que pretendam desenhar acções de formação para indivíduos mais velhos.

O género e as habilitações literárias têm alguma *influência conjunta* sobre as percepções da estrutura do módulo e da adequação do ensino. Muito embora as habilitações literárias por si só quase não tenham efeito sobre as percepções das mulheres sobre a estrutura do módulo, os homens com pelo menos o 10º ano fizeram uma avaliação significativamente menos positiva do que os menos instruídos. Os homens com mais habilitações também perceberam a estrutura do módulo significativamente de forma menos positiva que as mulheres com mais habilitações. Dentro do grupo de formandos com o 10º ano ou mais, as mulheres perceberam a adequação do ensino significativamente de forma mais positiva que os homens mas, dentro do grupo de formandos com escolaridade até ao 9º ano, as percepções dos homens e das mulheres não diferiram significativamente.

*A influência do curso de qualificação inicial sobre as percepções do Módulo.*

Os formandos que frequentavam cursos tradicionalmente masculinos menos exigentes (serralharia civil/mecânica and metalomecânica) deram mais valor à adequação do monitor que os que frequentavam cursos tradicionalmente mais femininos mais exigentes (esteticismo, secretariado comercial e práticas administrativas). Talvez possamos especular que provavelmente esta diferença de percepções não se deve ao curso que o formando está a frequentar mas antes a diferenças no género e no nível de habilitação dos formandos.

Os formandos do grupo misto formado por cursos de colarinho branco e cursos de colarinho azul avaliaram a estrutura e o valor do módulo menos positivamente que os dos outros três grupos de cursos (tradicionalmente femininos mais/menos exigentes e tradicionalmente masculinos menos exigentes). O porquê destas diferenças não é totalmente claro. Mas os formandos nos três grupos de cursos «tradicionalistas» tinham maiores expectativas do que os do grupo misto em termos de «melhoria do ambiente de trabalho» e «melhor relacionamento interpessoal» Deste modo, é possível que os formandos desses cursos percepcionassem a estrutura e o valor do módulo como mais relevante para estas expectativas do que os formandos do grupo misto.

Tanto os formandos deste grupo mais heterogéneo como os dos cursos tradicionalmente femininos menos exigentes (cabeleireiro e manicura/pedicura) perceberam de modo menos positivo a adequação do ensino. No entanto, nem todas as percepções dos formandos do grupo heterogéneo foram negativas. Tanto os formandos deste grupo como os dos dois grupos formados por cursos tradicionalmente frequentados por mulheres avaliaram mais positivamente a adequação da organização do módulo do que os homens do grupo de cursos tradicionalmente masculinos menos exigentes (serralharia civil, serralharia mecânica e metalomecânica).

Não é claro o porquê deste resultado, mas quase 60% dos formandos no grupo de cursos tradicionalmente frequentado por homens esperavam que o módulo lhes facultasse «melhor relacionamento interpessoal» (comparado com 35% ou menos de formandos nos outros três grupos). É pois possível que o grupo tradicionalmente masculino percepcionasse de forma mais negativa a organização do módulo pois não preencheu esta expectativa.

Contudo, um exame das médias dos três itens (Cumprimento o programa, linguagem utilizada and nível de participação permitido) que definiam o factor «adequação da organização» sugeriu outra explicação. Os formandos no grupo de cursos tradicionalmente frequentados por homens tinham as média mais baixa de todas apenas no item «cumprimento do programa». Assim, parece provável que a percepção mais negativa sobre a adequação da organização do módulo feita pelos formandos deste grupo se deva à sua avaliação mais negativa do modo como o módulo deu cumprimento ao programa.

#### *Relação entre percepções e a avaliação global sobre o Módulo*

Numa avaliação global feita usando uma única escala ordinal com quatro níveis, 52,3% dos formandos classificaram o módulo como «muito bom» ou «bom», 22% classificaram-no como «médio», 2,1% como «insuficiente» e 23,6% não responderam à questão. A avaliação global feita pelos formandos apenas está significativamente relacionada com as percepções sobre a estrutura e valor do módulo. Em conjunto, estas apresentaram um coeficiente de correlação múltipla de 0,404 ( $P < 0,0001$ ) com a avaliação global. Muito embora quase um quarto dos formandos não tenha feito uma avaliação global, não acreditamos que isso menospreze a validade das conclusões a que chegámos. Isto porque estes formandos eram muito semelhantes, em termos das suas características pessoais (género, idade e escolaridade) e das suas percepções (*score* médio no factores sobre a estrutura e valor do módulo), aos que atribuíram uma classificação «insuficiente» aquando da avaliação global do módulo. Também é interessante notar que 64,4% dos formandos que não avaliaram o módulo na globalidade pertenciam ao grupo misto de cursos de colarinho branco e cursos de colarinho azul, e que destes, 74,1% eram homens e 60% tinham menos que o 10º ano de escolaridade. Se estes formandos que não fizeram uma avaliação global do módulo fossem considerados como tendo classificado (para si) o módulo como «insuficiente», a correlação múltipla entre a avaliação global e as percepções sobre a estrutura e o valor do módulo decresce para 0,350 embora continue a ser altamente significativa ( $P < 0,0001$ ).

A estrutura do módulo é definida principalmente pela utilidade dos temas, aplicabilidade dos temas na função, articulação teoria/prática, adequação dos exercícios realizados e métodos e meios audiovisuais utilizados. Percepções da estrutura mais positivas indicam avaliações mais positivas destes aspectos do módulo. O valor do módulo é definido principalmente por aquisição de novos conhecimentos, métodos e técnicas, interesse dos temas, utilidade dos temas, aplicabilidade dos temas na função e coordenação do curso. Percepções do valor do módulo mais positivas indicam avaliações mais positivas destes aspectos do módulo. Quase todos estes itens dizem respeito a assuntos práticos e, tendo por base a sua natureza, podemos dizer que as avaliações globais que os formandos fizeram basearam-se principalmente nas suas percepções sobre a **relevância prática** do módulo.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ALMEIDA, J. F., FERNANDES, A.T. e Hill, M. M. (2000), Relatório PRAXIS sobre Competitividade e exclusão Social nas áreas metropolitanas de Lisboa e Porto, Dinâmia, Cies, FLUP.
- BARBIER, J. M. (1985), *A Avaliação em Formação*, Biblioteca das Ciências do Homem, Edições Afrontamento, nº 8.
- BRINKERHOFF, R. O. (1987), *Achieving Results from Training*, San Francisco, Jossey-Bass.
- EASTERBY-SMITH, M. (1994), *Evaluating Management Development, Training and Education*, 2<sup>nd</sup> Edition, Hampshire: Gower Publishing.
- HARRISON R. (1988), *Taining and Development*, London, Institute of Personnel Management.
- KIRKPATRICK, D. L. (1976), *Evaluation of training*, New York: Mc Graw - Hill.
- MEIGNANT, A. (1999), *A Gestão da formação*, Lisboa: Publicações Dom Quixote.
- NEWBY, A. C. (1992), *Training Evaluation Handbook*, Hampshire: Gower Publishing.

# A M | G | $\infty$ Systems Collection with Exponential Busy Period

Autor:  
Manuel Alberto M. Ferreira

**VOLUME II**

**2º QUADRIMESTRE DE 2002**

---

---

A  $M|G|\infty$  SYSTEMS COLLECTION WITH EXPONENTIAL BUSY PERIOD

---

---

---

---

UMA COLEÇÃO DE SISTEMAS  $M|G|\infty$  COM PERÍODO DE OCUPAÇÃO EXPONENCIAL

---

---

Autor: Manuel Alberto M. Ferreira  
- Professor Associado – Departamento de Métodos Quantitativos –  
I.S.C.T.E.

**ABSTRACT:**

- This paper describes how studying the transient behaviour of the  $M|G|\infty$  system, starting at the beginning of a busy period, it is possible to determine a collection of service distributions for which the length of the busy period is exponential, with an atom at the origin. We determine also the distribution of the busy cycle length for these systems and the mean number of busy periods that begin in  $[0, t]$ .

**KEY-WORDS:**

- $M|G|\infty$ , Busy Period, Busy Cycle, Ricatti's Equation, Exponential Distribution.

**RESUMO:**

- Este trabalho descreve como o estudo do comportamento transeunte do sistema  $M|G|\infty$ , sendo o instante inicial o do começo de um período de ocupação, permite determinar uma coleção de distribuições de serviço para as quais o comprimento do período de ocupação é exponencial, com uma concentração de probabilidade na origem. Determina-se também a distribuição do comprimento do ciclo de ocupação para estes sistemas e o número médio de períodos de ocupação que se iniciam em  $[0, t]$ .

**PALAVRAS-CHAVE:**

- $M|G|\infty$ , Período de Ocupação, Ciclo de Ocupação, Equação de Ricatti, Distribuição Exponencial.

**VOLUME II**

**2º QUADRIMESTRE DE 2002**

## 1. INTRODUCTION

The determination of the busy period length distribution is one of the most delicate subjects in queues studies. Here we present a service distributions collection for which the  $M|G|\infty$  queue busy period has exponential distribution, with an atom at the origin.

In a  $M|G|\infty$  queue the arrival process is Poisson, at rate  $\lambda$ , and every customer meets, upon its arrival, a server available and receives a service whose length is a positive random variable with distribution function  $G(\cdot)$ . We will call  $i(t)$  the integral between 0 and  $t$  of  $1 - G(v)$ , being  $v$  the variable, and we have  $i(\infty) = \alpha$ , being  $\alpha$  the mean service time. The traffic intensity,  $\rho$ , is given for  $\rho = \lambda\alpha$ .

A queue system busy period begins when a customer arrives at the system finding it empty, ends when a customer leaves it letting it empty, and during it there is always at least one customer in the queue.

So, in a queue system, there is a sequence of busy and idle periods. A busy period followed by an idle period is called a busy cycle.

Calling  $p_{10}(t)$  the  $M|G|\infty$  system emptiness probability at time  $t$ , being the time origin the beginning of a busy period, and  $p_{00}(t)$  the same probability in the situation of initially empty system, we have

$$p_{10}(t) = G(t)p_{00}(t) \quad (1.1).$$

In fact, so that the system is empty at time  $t$ , the customer that arrived in the initial instant must have left, what happens with probability  $G(t)$ , and the servers that were unoccupied at time origin must go on unoccupied at time  $t$ , what happens with probability  $p_{00}(t)$ . The independence of these two facts justifies (1.1).

As  $p_{00}(t) = \exp[-\lambda i(t)]$  (Harrison and Lemoine, 1981), from (1.1), we have

$$p_{10}(t) = G(t)\exp[-\lambda i(t)] \quad (1.2).$$

---

2. DETERMINATION OF THE SERVICE DISTRIBUTIONS COLLECTION FOR WHICH THE BUSY PERIODS  $M|G|\infty$  QUEUE IS EXPONENTIALLY DISTRIBUTED

---

The  $M|G|\infty$  queue busy period length Laplace-Stieltjes transform,  $\bar{B}(s)$ , is given by

$$\bar{B}(s) = \frac{p_{10}(s)}{p_{00}(s)} \quad (2.1)$$

where  $p_{10}(s)$  and  $p_{00}(s)$  are the Laplace-Stieltjes transforms of  $p_{10}(t)$  and  $p_{00}(t)$ , respectively (Ferreira, 1995).

From (1.2)  $\frac{d}{dt} p_{10}(t) = p_{00}(t)(1 - G(t)) \left( \frac{g(t)}{1 - G(t)} - \lambda G(t) \right)$ . Putting

$\frac{g(t)}{1 - G(t)} - \lambda G(t) = \beta$ ,  $t > 0$  with  $\beta$  real, we get the differential equation

$$\frac{dG(t)}{dt} = -\lambda G^2(t) + (\lambda - \beta)G(t) + \beta \quad (2.2)$$

that is a Riccati equation (see, for instance, (Ferreira, 1994)). To solve it we have to put  $G(t) = u(t) + \frac{1}{z(t)}$  where  $u(t)$  is a solution of (2.2). It is easily seen that we can have  $u(t) = 1$  getting

$$\frac{dz}{dt} - (\lambda + \beta)z = \lambda \quad (2.3)$$

The equation (2.3) is a first order linear differential equation with constant coefficients easily solvable. So we get

$$G(t) = 1 - \frac{(1 - e^{-\rho})(\lambda + \beta)}{\lambda e^{-\rho}(e^{(\lambda + \beta)t} - 1) + \lambda}, \quad t \geq 0, \quad -\lambda \leq \beta \leq \frac{\lambda}{e^{\rho} - 1} \quad (2.4)$$

that is the required distribution functions collection. In fact putting (2.4) at (2.1) we get

$$\bar{B}(s) = \frac{e^{-\rho}(\lambda + \beta)(\lambda + s) - s\beta}{\lambda(e^{-\rho}(\lambda + \beta) + s)} \quad (2.5)$$

whose inversion gives the p.d.f. for the  $M|G|\infty$  queue busy period length

$$b(t) = \frac{e^{-\rho}(\lambda + \beta) - \beta}{\lambda} \delta(t) + \left(1 - \frac{e^{-\rho}(\lambda + \beta) - \beta}{\lambda}\right) e^{-\rho}(\lambda + \beta) e^{-e^{-\rho}(\lambda + \beta)t},$$

$$t \geq 0, -\lambda \leq \beta \leq \frac{\lambda}{e^{\rho} - 1} \quad (2.6),$$

being  $\delta(\cdot)$  the Dirac function.

Note that:

$$- \beta = -\lambda$$

Then

$$G(t) = 1, \quad t \geq 0 \quad \text{and} \quad b(t) = \delta(t) \quad (2.7)$$

(see (Ferreira , 1988)).

$$- \beta = 0$$

Then

$$G(t) = \frac{e^{-\rho}}{e^{-\rho} + (1 - e^{-\rho})e^{-\lambda t}}, \quad t \geq 0 \quad (2.8)$$

and

$$b(t) = e^{-\rho} \delta(t) + (1 - e^{-\rho}) e^{-\rho} \lambda e^{-e^{-\rho} \lambda t}, \quad t \geq 0 \quad (2.9)$$

(see (Ferreira , 1991)).

$$- \beta = \frac{\lambda}{e^{\rho} - 1}$$

Then

$$1 - G(t) = \frac{1}{1 - e^{-\rho} + e^{-\rho + \frac{\lambda}{1 - e^{-\rho}} t}}, \quad t \geq 0 \quad (2.10)$$

and

$$b(t) = \frac{\lambda}{e^{\rho} - 1} e^{-\frac{\lambda}{e^{\rho} - 1} t}, \quad t \geq 0 \quad (2.11)$$

(see (Ferreira, 1995b)).

The distribution function corresponding to (2.6) is

$$B(t) = 1 - \frac{\lambda + \beta}{\lambda} (1 - e^{-\rho}) e^{-e^{-\rho}(\lambda + \beta)t}, \quad t \geq 0, \quad -\lambda \leq \beta \leq \frac{\lambda}{e^{\rho} - 1} \quad (2.12).$$

If  $0 \leq \beta \leq \frac{\lambda}{e^{\rho} - 1}$ , for  $\alpha$  great enough so that  $1 - e^{-\rho} \cong 1$ ,  $G(t) \cong 0$  and  $B(t) \cong 1 - e^{-e^{-\rho}\lambda t}$  (see (Ferreira e Ramalhoto, 1994)).

If  $-\lambda < \beta < 0$ , for  $\alpha$  great enough so that  $1 - e^{-\rho} \cong 1$ ,  $G(t) \cong -\frac{\beta}{\lambda}$  and  $B(t) \cong 1 - \frac{\lambda + \beta}{\lambda} e^{-e^{-\rho}(\lambda + \beta)t}$ . So, only for  $\lambda$  great enough we have  $G(t) \cong 0$  and  $B(t) \cong 1 - e^{-e^{-\rho}\lambda t}$  (see (Ferreira, 1996)).

That is, the busy period, for these systems, shows exponential behaviour in “heavy-traffic” situations.

### 3. BUSY CYCLE

The  $M|G|\infty$  queue busy length Laplace-Stieltjes transform is given by

$$\bar{Z}(s) = \frac{\lambda}{\lambda + s} \bar{B}(s) \quad (3.1)$$

because the idle period length is exponential with parameter  $\lambda$  (as it happens with any queue systems with Poisson arrival process) and is independent of the busy period length (Takács, 1962).

Substituting (2.5) in (3.1) and inverting we get the  $M|G|\infty$  queue busy cycle length p.d.f.

$$z(t) = \frac{e^{-\rho}(1 - e^{-\rho})(\lambda + \beta)^2}{\lambda - e^{-\rho}(\lambda + \beta)} e^{-e^{-\rho}(\lambda + \beta)t} - \frac{\lambda\beta}{\lambda - e^{-\rho}(\lambda + \beta)} e^{-\lambda t},$$

$$t \geq 0, -\lambda \leq \beta \leq \frac{\lambda}{e^{\rho} - 1} \quad (3.2).$$

So,

$$- \beta = -\lambda$$

$$z(t) = \lambda e^{-\lambda t}, \quad t \geq 0, \quad (3.3)$$

evidently,

$$- \beta = 0$$

$$z(t) = \lambda e^{-\rho} e^{-\lambda e^{-\rho} t}, \quad t \geq 0, \quad (3.4)$$

(see (Ferreira, 1995a)).

$$- \beta = \frac{\lambda}{e^{\rho} - 1}$$

$$z(t) = \frac{\lambda}{e^{\rho} - 2} \left( e^{-\frac{\lambda}{e^{\rho} - 1} t} - e^{-\lambda t} \right), \quad t \geq 0, \quad (3.5).$$

---

4. MEAN NUMBER OF BUSY PERIODS THAT BEGIN IN  $[0, t]$

---

Be the  $M|G|\infty$  queue with time origin at the beginning of a busy period. The instants  $0, t_1, t_2, \dots$  at which begin a busy period, are the arrival moments of a renewal process (Takács, 1962). This renewal process function is (Ferreira, 1995)

$$R(t) = e^{-\lambda \int_0^t [1-G(v)]dv} + \lambda \int_0^t e^{-\lambda \int_0^u [1-G(v)]dv} du \quad (4.1)$$

and gives the mean number of busy periods that begin in  $[0, t]$ .

Substituting (2.5) in (4.1) we get

$$R(t) = 1 + \lambda t, \quad \beta = -\lambda \quad (4.2)$$

according to (3.3), and

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

$$-\lambda < \beta \leq \frac{\lambda}{e^{\rho} - 1} \quad (4.3).$$

So,

$$-\beta = 0$$

$$R(t) = 1 + \lambda e^{-\rho} t \quad (4.4)$$

according to (3.4),

$$-\beta = \frac{\lambda}{e^{\rho} - 1}$$

$$R(t) = e^{-\rho} + (1 - e^{-\rho})^2 + \lambda e^{-\rho} t + e^{-\rho} (1 - e^{-\rho}) e^{-\frac{\lambda}{1 - e^{-\rho}} t} \quad (4.5)$$

(see, (Ferreira, 1995)).

## 5. CONCLUSIONS

Imposing  $\frac{g(t)}{1-G(t)} - \lambda G(t)$  to be constant, we find a service distributions family for which the  $M|G|\infty$  queue busy period length has a distribution that is a mixture of an origin degenerated distribution and an exponential one.

Other distributions already determined are members of this family for which the busy period length has a similar behaviour.

We have, so, a  $M|G|\infty$  systems collection for which everything is known about the busy period distribution and about the busy cycle distribution.

It happens also that  $b(t)$  and  $z(t)$ , for these  $M|G|\infty$  systems collection, have quite simple structures that do not rise any problem as for the distribution function computation and as for the moments of any order.

Finally, for such an important parameter as the mean number of busy periods that begin in  $[0, t]$ , we get also quite simple expressions for these systems.

## REFERENCES

- FERREIRA, M.A.M., (1988) "Redes de Filas de Espera". Dissertação de Mestrado discutida no I.S.T.. Lisboa.
- FERREIRA, M.A.M., (1991) "Um Sistema  $M|G|\infty$  com período de ocupação exponencial". Actas das XV Jornadas Luso-Espanholas de Matemática. Vol. IV. Universidade de Évora. Évora.
- FERREIRA, M.A.M., (1994) "Matemática – Integrais Múltiplos e Equações Diferenciais". Edições Sílabo. Lisboa. 4.ª Edição.
- FERREIRA, M.A.M., (1995 a) "Comportamento Transeunte e Período de Ocupação de Sistemas de Filas de Espera Sem Espera". Dissertação de Doutoramento discutida no I.S.C.T.E.. Lisboa.
- FERREIRA, M.A.M., (1995 b) "Cauda do Período de Ocupação da Fila de Espera  $M|G|\infty$ ". Bom Senso e Sensibilidade – Traves Mestras da Estatística. Actas do III Congresso Anual da S.P.E.. Guimarães.
- FERREIRA, M.A.M., (1996) "Distribuição do Comprimento do Período de Ocupação da Fila de Espera  $M|G|\infty$  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. e RAMALHOTO, M.F., (1994) “Estudo dos Parâmetros Básicos do Período de Ocupação da Fila de Espera  $M|G|\infty$ ”. A Estatística e o Futuro da Estatística. Actas do I Congresso Anual da S.P.E.. Edição Salamandra. Lisboa.

HARRISON, J.M. e LEMOINE, A.J., (1981) “ A Note on Networks of Infinite Server Queues”, J.A.P. 18, 561-567.

TACKÁS, L., (1962) “An Introduction To Queueing Theory”. Oxford University Press. New York.

# INFORMAÇÕES

**VOLUME II**

**2° QUADRIMESTRE DE 2002**

---

---

CONGRESSOS, SEMINÁRIOS, COLÓQUIOS E CONFERÊNCIAS

---

---

CONGRESS, SEMINARS AND CONFERENCES

---

---

**2002**

- 15-19 September  
**Annual Meeting of the International Association for Mathematical Geology, IAMG 2002**, to be held at the Freie Universität Berlin, Germany. Email: [iang2002@zedat.fu-berlin.de](mailto:iang2002@zedat.fu-berlin.de) Website: <http://www.fu-berlin.de/iang2002/>  
Information: E-mail: [iang2002@zedat.fu-berlin.de](mailto:iang2002@zedat.fu-berlin.de)  
URL: <http://www.fu-berlin.de/iang2002/>
  
- 16-18 September  
**KOMAT'2002: International Workshop on Intelligent Knowledge Management Techniques (I-KOMAT'2002)** to be held in conjunction with: KES'2002 (Sixth International Conference on Knowledge-Based Intelligent Information & Engineering Systems), Podere d'Ombriano, Crema, Italy.  
Information: [www-mugc.cc.monash.edu.au/~maum/ikomat-02.htm](http://www-mugc.cc.monash.edu.au/~maum/ikomat-02.htm)  
Information: URL: [www-mugc.cc.monash.edu.au/~maum/ikomat-02.htm](http://www-mugc.cc.monash.edu.au/~maum/ikomat-02.htm)
  
- 20-22 September  
**2nd Conference in Actuarial Science and Finance**, to be held in Samos, Greece.  
Information: Website: [www.stat.ucl.ac.be/Samos2002/](http://www.stat.ucl.ac.be/Samos2002/)  
Information: URL: [www.stat.ucl.ac.be/Samos2002/](http://www.stat.ucl.ac.be/Samos2002/)
  
- 20-25 September  
**"The Humanistic Renaissance in Mathematics Education"**, to be held at the Citta del Mare Hotel, near the city of Palermo, Sicily, Italy.  
Email [arogerson@vsg.edu.au](mailto:arogerson@vsg.edu.au)  
Information: E-mail: [arogerson@vsg.edu.au](mailto:arogerson@vsg.edu.au)
  
- 24-27 September  
**Third International ICSC Symposium on Engineering of Intelligent Systems and the ISMC 2002 Workshop Workshop on Information Systems for Mass Customization**, to be held at the University of Malaga, Malaga, Spain.  
  
Information: ICSC- NAISO The Netherlands (Operating Division), P.O. Box 1091, 3360 BB Sliedrecht, The Netherlands. Phone: +31-184-496999, Fax: +31-184-421065.  
E-mail: [eis2002@ITStransnational.com](mailto:eis2002@ITStransnational.com)  
[planning@icsc.ab.ca](mailto:planning@icsc.ab.ca)  
URL: [www.icsc-naiso.org/conferences/eis2002/index.html](http://www.icsc-naiso.org/conferences/eis2002/index.html)

- 26-27 September  
**"Innovative Solutions in providing access to micro-data"**, seminar organised by the CEIES subcommittee on Dissemination Policy, in co-operation with the Instituto Nacional de Estatística, to be held in Lisbon.  
Information: Ms Annika Näslund-Fogelberg, CEIES Secretariat, Eurostat, +352 4301 33055; Ms Nicole Lauwerijs, assistant, Eurostat, +352 4301 33564  
E-mail: [annika.naslund-fogelberg@cec.eu.int](mailto:annika.naslund-fogelberg@cec.eu.int)  
[nicole.lauwerijs@cec.eu.int](mailto:nicole.lauwerijs@cec.eu.int)  
URL: <http://forum.europa.eu.int/Public/irc/dsis/ceies/library>
  
- 25-28 September  
**The 10th Annual Conference of the Portuguese Statistical Society (SPE)** will be held in Porto, Portugal.  
Information: E-mail: [spe2002@fep.up.pt](mailto:spe2002@fep.up.pt)  
URL: <http://www.fep.up.pt/spe2002>
  
- 16-19 October  
**26th CIRET Conference on "Business Surveys, Business Cycle Indicators and Consumer Surveys"**, to be held in Taipei/Taiwan, China.  
Information: URL: [www.ciret.org/](http://www.ciret.org/),
  
- 17-18 October  
**Third francophone Conference on sampling**, in Grenoble and Autrans (France).  
Information: Benoît Riandey  
E-mail: [riandey@ined.fr](mailto:riandey@ined.fr)  
URL: [sondages2002.upmf-grenoble.fr/](http://sondages2002.upmf-grenoble.fr/)
  
- 17-18 October  
**The Second Thiele Symposium on Financial Econometrics**, University of Copenhagen, Denmark.  
Information: David Lando, Thomas Mikosch and Michael Sørensen.  
URL: <http://www.math.ku.dk/~michael/thiele2/>
  
- 21-23 October  
**International clinical trials symposium**, to be held in Sydney, Australia.  
Information: E-mail: [trials@icms.com.au](mailto:trials@icms.com.au)  
URL: [www.ctc.usyd.edu.au/4news/Symposium2002/ICTS\\_home.htm](http://www.ctc.usyd.edu.au/4news/Symposium2002/ICTS_home.htm)
  
- 22-23 October  
**International Conference on Quality and Innovation**, to be held at the Waterloo Inn, Waterloo, Ontario, Canada.  
Information: Institute for Improvement in Quality and Productivity and the American Society for Quality (ASQ) - Quality Management Division (QMD).  
URL: <http://www.iiqp.uwaterloo.ca/Events/qi.html>

- 31 October-03 November  
**Euroworkshop on Statistical Modelling - Model Building and Evaluation**, to take place in Schloss Hoehenried, Bernried, Munich, Germany.  
Information: Herwig Friedl.  
 E-mail: [friedl@stat.tu-graz.ac.at](mailto:friedl@stat.tu-graz.ac.at)  
           [goeran@stats.gla.ac.uk](mailto:goeran@stats.gla.ac.uk)  
 URL: <http://www.stat.uni-muenchen.de/euroworkshop/2002.html>
  
- 06-08 November  
**XIX International Methodology Symposium "Modelling Survey Data for Social and Economic Research"** to be held in Ottawa, Canada.  
Information: E-mail: [Symposium2002@statcan.ca](mailto:Symposium2002@statcan.ca)  
 URL: <http://www.statcan.ca/english/services/smnrs.htm>  
 URL: [http://www.statcan.ca/francais/services/smnrs\\_f.htm](http://www.statcan.ca/francais/services/smnrs_f.htm)
  
- 14-17 November  
**International Conference on Questionnaire Development, Evaluation, and Testing, Methods**, to be held in Charleston, South Carolina, USA.  
Information: Jennifer Rothgeb  
 E-mail: [jennifer.m.rothgeb@census.gov](mailto:jennifer.m.rothgeb@census.gov)  
 or  
 URL: <http://www.jpsm.umd.edu/qdet>
  
- 20-23 November  
**International Conference on Information and Communication Technologies in Education**, to be held at the Universidad de Extremadura, Badajoz, Spain.  
Information: Antonio Mendez Vilas.  
 E-mail: [amvilas@unex.es](mailto:amvilas@unex.es)  
 URL: <http://www.formatex.org/congreso.html>
  
- 25-27 November  
**"Integration of Statistics in the Information Society"**, Hotel Radin , Radenci , Slovenia.  
Information: E-mail: [andreja.hocevar@gov.si](mailto:andreja.hocevar@gov.si)  
 URL: <http://www.sigov.si/zrs/eng/rad02/aradenci.htm>
  
- 05-07 December  
**4th Conference of the Asian Regional Section of the IASC. "Theme: e-Statistics for Information Society"**, The Westin Chosun Beach Hotel, Busan, Korea.  
Information: URL: <http://www.chosunbeach.co.kr/>
  
- 09-13 December  
**9th Annual Biopharmaceutical Applied Statistics Symposium**, Savannah, Georgia, United States.  
Information: MCV/VCU & GASOU Biostatistics Depts.  
 E-mail: [KEPeace@gasou.edu](mailto:KEPeace@gasou.edu)  
 URL: <http://views.vcu.edu/bis/bass/>

- 21-24 December

**Eighth Islamic Countries Conference on Statistical Science**, to be organised by the Islamic Society of Statistical Sciences (ISOSS) in collaboration with the University of Bahrain. The conference will take place at the University of Bahrain.

Information: Organizing Committee, College of Business Administration, Department of Management & Marketing, University of Bahrain, P.O. Box # 32038, Sakhir, Bahrain. Tel: +973-449408/449413/449400, Fax: +973-449476, OR Mr. Akhlaq Ahmad, Executive Director, Islamic Society of Statistical Sciences, RB/2, 2nd Floor, Awami Complex, Usman Block, New Garden Town, Lahore-54600 (Pakistan).  
E-mail: <mailto:akrammoh@internic.uob.bh>  
[drmunir@brain.net.pk](mailto:drmunir@brain.net.pk)

- 28-30 December

**International Conference on "Ranking and Selection, Multiple Comparisons, Reliability, and Their Applications"**. Tentative Venue: Hotel Savera, Chennai, Tamilnadu, India.

**Organizers:** [bala@mcmail.cis.mcmaster.ca](mailto:bala@mcmail.cis.mcmaster.ca); [NKannan@utsa.edu](mailto:NKannan@utsa.edu); H. N. Nagaraja, Ohio State University, <mailto:hnn@stat.ohio-state.edu>

Information: N. Balakrishnan, McMaster University; N. Kannan, University of Texas at San Antonio; H. N. Nagaraja, Ohio State University.

E-mail: [bala@mcmail.cis.mcmaster.ca](mailto:bala@mcmail.cis.mcmaster.ca)  
[NKannan@utsa.edu](mailto:NKannan@utsa.edu)  
<mailto:hnn@stat.ohio-state.edu>

## **2003**

- 02-04 January

**International Conference on 'Statistics in industry and business'** to be held in Cochin, India.

Information: Bovas Abraham, IIQP Univ of Waterloo, Waterloo, Ontario N2L 3G1, Canada; Tele: 519-888-4593, Fax:519-746-5524; Dr.P.G.Sankaran Dept. of Statistics Cochin University of Science & Technology Cochin-682022; Tel.no: 0484-555893(Office), 0484-741693(Res.).

E-mail: [babraham@uwaterloo.ca](mailto:babraham@uwaterloo.ca)  
[pgsankaran@cusat.ac.in](mailto:pgsankaran@cusat.ac.in),  
[sankaranpg@yahoo.com](mailto:sankaranpg@yahoo.com)

- 08-10 January

**The seventh conference on "Statistics and youth assistance in big cities", the forth SCORUS Berlin meeting** to be held in "Haus am Rupenhorn".

Information: Prof. Dr. Eckart Elsner; Statistisches Landesamt Berlin; 10306 Berlin; Tel. +49 (30) 90 21 34 88; Fax +49 (30) 90 21 25 61

E-mail: [e.elsner@statistik-berlin.de](mailto:e.elsner@statistik-berlin.de)  
URL: <http://www.statistik-berlin.de/>

- 12-15 January  
**2nd International Conference of EMR-IBS. The 2nd International Conference in Biostatistics of the Eastern Mediterranean Region of the International Biometric Society** will be held at the BelConti Resort Hotel in Antalya, Turkey.  
Information: URL: <http://biostat.cu.edu.tr/emr>
  
- 08-11 June  
**Annual Meeting of the Statistical Society of Canada**, Halifax, Nova Scotia, Canada.  
Information: Chris Field Department of Mathematics and Statistics, Dalhousie University, Halifax, Nova Scotia, B3H 3J5, Canada; Phone: (902) 494-3339 or (902) 494-2572 Fax: (902) 494-5130.  
E-mail: [field@mscs.dal.ca](mailto:field@mscs.dal.ca)
  
- 09-12 June  
**International Conference on "Advances in Statistical Inferential Methods"** (ASIM) Almaty, Republic of Kazakhstan.  
Information: E-mail: [voinov@kimep.kz](mailto:voinov@kimep.kz)  
URL: [www.kimep.kz/research/asim](http://www.kimep.kz/research/asim)
  
- 24-27 June  
**International NAISO Symposium on Information Technologies in Environmental Engineering (ITEE'2003)**, to be held at the Technical University of Gdansk, Poland.  
Information: E-mail: [itee2003@global-conf.org](mailto:itee2003@global-conf.org)  
URL: <http://www.icsc-naiso.org/conferences/itee2003/index.html>
  
- 24-27 June  
**First World Congress on Information Technology in Environmental Engineering ITEE 2003**, Technical University of Gdansk, Poland.  
Information: URL: <http://www.icsc-naiso.org/conferences/itee2003/index.html>
  
- 29 June-02 July  
**International Conference on Correspondence Analysis and Related Methods (CARME 2003)**, to be held at the Universitat Pompeu Fabra, Barcelona, Catalonia, Spain.  
Information: E-mail: [carme2003@upf.es](mailto:carme2003@upf.es)  
URL: <http://www.econ.upf.es/carme>
  
- 14-18 July  
**International Conference on Robust Statistics 2003 (ICORS 2003)**, to be held at the University of Antwerp, Belgium.  
Information: Stefan Van Aelst  
E-mail: [statis@uia.ua.ac.be](mailto:statis@uia.ua.ac.be)  
URL: <http://wia-www.uia.ac.be/u/statis/Icors03.htm>

- 20-24 July  
**ISCB 24 - 24th Meeting of the International Society for Clinical Biostatistics**, to be held in London, UK.  
Information: E-mail: [diana.elbourne@lshtm.ac.uk](mailto:diana.elbourne@lshtm.ac.uk)  
 URL: <http://www.iscb-homepage.org/>
- 05-08 August  
**12th IWMS - International Workshop on Matrices and Statistics**, to be held at the University of Dortmund, Department of Statistics, Dortmund, Germany.  
Information: E-mail: [iwms2003@statistik.uni-dortmund.de](mailto:iwms2003@statistik.uni-dortmund.de)  
 URL: <http://www.statistik.uni-dortmund.de/IWMS/main.html>
- 13-20 August  
**International Statistical Institute, 54th Biennial Session** (includes meetings of the Bernoulli Society, The Intern. Assoc. for Statistical Computing, The Intern. Assoc. of Survey Statisticians, The Intern. Assoc. for Official Statistics and The Interna. Assoc. 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](mailto:isi@cbs.nl)  
 or visit the Session website at <http://www.isi-2003.de>
- 10-12 December  
**World Summit on Information Society**, to be held in Geneva, Switzerland.  
Information: Alain Clerc, Director, Civil-Society Division - Executive Secretariat,  
 World Summit on the Information Society.  
 E-mail: [info@geneva2003.org](mailto:info@geneva2003.org)  
 URL: <http://www.geneva2003.org/>

## **2004**

- 30 May-02 June  
**Annual Meeting of the Statistical Society of Canada**, Montreal, Quebec, Canada.  
Information: Christian Léger, Département de mathématiques et de statistique,  
 Université de Montréal. C.P. 6128, succursale Centre-ville. Mon-  
 tréal, Québec, Canada H3C 3J7; Tel.: (514) 343-7824; Fax: (514)  
 343-5700  
 E-mail: [leger@dms.umontreal.ca](mailto:leger@dms.umontreal.ca)
- 04-11 July  
**10th International Conference on Mathematical Education (ICME 10)**, to be held in Copenhagen, Denmark.  
Information: URL: <http://www.icme-10.dk/>

- 20-24 July  
**ISCB 25 - 25th Meeting of the International Society for Clinical Biostatistics**,  
to be held in Leiden, Netherlands.  
Information: E-mail: [jcvanhouwelingen@lumc.nl](mailto:jcvanhouwelingen@lumc.nl)  
URL: <http://www.iscb-homepage.org/>

## **2005**

- 05-12 April  
**International Statistical Institute, 55th 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 Sydney, Australia.  
Information: 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](mailto:isi@cbs.nl)

**VOLUME II**

**2º QUADRIMESTRE DE 2002**

## FUNDAMENTO, OBJECTO E ÂMBITO DA REVISTA

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, passou 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 do Sistema Estatístico Nacional;
- c) Informações sobre acções desenvolvidas pelo INE no âmbito da cooperação.
- d) Informações sobre congressos, seminários, colóquios e conferências de interesse estatístico ou afim;

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 de artigos por iniciativa dos próprios autores serão objecto de avaliação de mérito científico pelo Conselho Editorial, que decidirá ou não pela sua publicação.

Para a elaboração e envio das contribuições de artigos para publicação na Revista são adoptadas as *Normas de Apresentação de Originais* 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 OF THE REVIEW

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 of the National Statistical System;
- c) Information on activities developed by INE within the scope of co-operation;
- d) Information on congresses, seminars and conferences of a statistical or related nature;

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 Originals* 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 ORIGINAIS

Nos termos do *Regulamento da Revista de Estatística*, o Conselho Editorial aprovou as seguintes **Normas de Apresentação de Originais**:

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 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 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 original apresentado.
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, com indicação do(s) software(s) adicional(ais) eventualmente utilizado(s) na produção do documento original, e entregues em suporte papel acompanhado da respectiva *disquette*, ou enviados por E-mail para o seguinte endereço: [liliana.martins@ine.pt](mailto:liliana.martins@ine.pt)
6. Na apresentação dos originais, os autores respeitarão ainda as seguintes normas:
  - 6.1. Quanto à *estrutura*:
    - a) O texto deve ser processado em formato *A<sub>4</sub>*, com utilização do tipo de letra *Times New Roman 11*, espaçamento *at least 12*, e com as seguintes margens: *top*: 4 cm, *bottom*: 3 cm, *left*: 2,5 cm, *right*: 5 cm, *header*: 1,25cm, *footer*: 1,25cm;
    - b) A primeira página conterà exclusivamente o título do artigo, bem como o nome, morada e telefone, fax e E-mail 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 100 palavras,

seguido de um parágrafo com indicação de *palavras-chave* até ao limite de 15;

- 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:

GREENE, W. H., “*Econometric Analysis*”, Prentice-Hall, New Jersey, 1993.

6.3. Quanto à *revisão de provas e publicação*:

- a) Uma vez aceite o artigo e antes da sua publicação, receberá o autor provas para revisão, as quais serão devolvidas 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 Director 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.<sup>a</sup>. António José de Almeida, n.º 5 – 9.º.  
1000-043 Lisboa - Portugal

- Tel.: +351 21 842 62 05
- Fax.: +351 21 842 63 84
- e-mail: [liliana.martins@ine.pt](mailto:liliana.martins@ine.pt)

## RULES FOR SUBMITTING ORIGINALS

Within the terms of the *Regulation of the Statistical Review*, the Editorial Committee has approved the following **Rules for Submitting Originals**:

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's 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 the articles submitted on the initiative of the 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 of scientific merit within a maximum period of 30 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, the original will be returned to the author.
4. The articles accepted for publication in the *Statistical Review* will also be made public on the Internet site of the INE.
5. The original articles having no more than thirty pages must be processed in *Word for Windows*, completely at black and white, with the information on the additional(s) software(s) eventually used in the production of the original document, and they will be delivered in hard copy as well as on diskette, or sent by E-mail to: [liliana.martins@ine.pt](mailto:liliana.martins@ine.pt)
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, spacing at least 12, and with the margins: *top* 4cm, *bottom* 3cm, *left* 2,5cm, *right* 5cm, *header* 1,25cm, *footer* 1,25cm;
    - b) The first page shall contain only the title of the article as well as the name, address and telephone, fax and E-mail 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 should be forwarded;

- c) The second page shall contain in *Portuguese* and *English* only the *title* and an *abstract* of the article with the maximum of 100 words followed by a paragraph indicating *key words* up to the limit of 15;
- 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:

GREENE, W. H., "*Econometric Analysis*", Prentice-Hall, New Jersey, 1993.

6.3 Regarding *proof-reading and publication*:

- a) Once the article is accepted and prior to its publication, the author will receive a copy for review. These copy 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 Director 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:

Eduarda Liliana Martins  
Instituto Nacional de Estatística  
Av<sup>a</sup>. António José de Almeida, n.º. 5 – 9.º.  
1000-043 Lisbon - Portugal

- ☐ Tel.: +351 1 21 842 62 05
- ☐ Fax.: +351 1 21 842 63 84
- ☐ e-mail: [liliana.martins@ine.pt](mailto:liliana.martins@ine.pt)