

Boletim



**SOCIEDADE PORTUGUESA
DE ESTATÍSTICA**

Publicação semestral

Primavera de 2009



Investigação (em) Estatística

CMA - Universidade Nova de Lisboa por C. A. Coelho, M. L. Esquível e J. T. Mexia	10
GPE - Universidade de Aveiro por Manuel Scotto	22
CIMA - Universidade de Évora por Carlos A. Braumann e Russell Alpizar-Jara	24
CEMAT - Universidade Técnica de Lisboa por António Pacheco	29
CEAUL - Universidade de Lisboa por Maria Isabel Fraga Alves	40
Algumas Reflexões Avulsas sobre a S.P.E por M. Ivette Gomes	50
Congressos da SPE: Um barómetro da investigação em Estatística realizada em Portugal por João A. Branco ..	53
Introduzindo <i>Outliers</i> por Fernando Rosado	60
Um olhar sobre a investigação estatística em Portugal por Carlos A. Braumann	65
Investigação (em) Estatística: da biologia à estatística ou perdido no limbo por T. A. Marques	67
Surfando pelos mares da biologia: peripécias de um jovem estatístico aspirante a biólogo por N. Sepúlveda ..	75

Editorial	1
Mensagem do Presidente	2
Notícias	3
Episódios na História da Estatística	7
SPE e a Comunidade	82
Ciência Estatística	
• Artigos Científicos Publicados	103
• Livros e Capítulos em Livros	104
• Teses de Mestrado	104
• Teses de Doutoramento	106
Prémios Estatístico Júnior	109

Informação Editorial

Endereço: Sociedade Portuguesa de Estatística.
Campo Grande. Bloco C6. Piso 4.
1749-016 Lisboa. Portugal.

Telefone: +351.217500120

e-mail: spe@fc.ul.pt

URL: <http://www.spestatistica.pt>

ISSN: 1646-5903

Depósito Legal: 249102/06

Tiragem: 1000 exemplares

Execução Gráfica e Impressão: Gráfica Sobreireense

Editor: Fernando Rosado, fernando.rosado@fc.ul.pt

Este Boletim tem o apoio da **FCT** Fundação para a Ciência e a Tecnologia

MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR



SOCIEDADE PORTUGUESA
DE ESTATÍSTICA



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

XVII CONGRESSO da SOCIEDADE PORTUGUESA DE ESTATÍSTICA

Sesimbra

30 de Setembro
a 3 de Outubro de 2009



Editorial

... “trintão”...

Este número do Boletim (também) celebra 30 anos de Investigação e de Estatística em Portugal. O Boletim SPE, como se sabe, nasceu a partir do Boletim Informativo de Estatística e Investigação Operacional e o Número 1 foi publicado em Fevereiro de 1979. Há 30 anos! Aliás, as convicções reveladas no Editorial dessa publicação histórica manifestam a gestação de uma nova sociedade científica. A SPEIO / SPE nasceu no ano seguinte. Assim, caminhando à frente, o Boletim foi (e ainda deve ser!) um arauto da Estatística.

Como homenagem, na secção Episódios na História da Estatística, apresentamos o *fac-símile* das suas oito páginas. Pelo esforço e a generosidade de muitos autores, esta publicação tem conseguido uma actividade regular ao longo de trinta anos e de que muito se deve orgulhar. Aos editores meus antecessores presto homenagem e gratidão pelo empenho e o zelo com que contribuíram para que o Boletim sempre tivesse conseguido avançar - muitas vezes em situações bastante difíceis. Como actual responsável pela edição renovo os meus agradecimentos a todos os autores que dedicadamente têm colaborado ajudando assim a criar novos exemplares.

Através das suas páginas, o *Boletim SPE*, além de informações específicas, divulga muita investigação que se faz em Estatística - em torno de temas centrais e assuntos de índole mais abrangente. A sua leitura, acima de tudo, deve proporcionar conhecimento do panorama científico no campo da Estatística em Portugal. Então, cada exemplar é dedicado à Investigação e à Estatística. A presente edição festiva, celebra e realça esse desiderato: A Investigação, a Estatística e também a Investigação (em) Estatística.

Como se faz a Investigação (em) Estatística? Quem faz? O que faz? Esta edição apresenta um breve ponto da situação.

O Boletim SPE “vive” da investigação e, assim, cada edição é um contributo para o melhor conhecimento de como e o que se está a fazer em determinado assunto. Nesta edição colocamos a Investigação no Tema Central. Importa saber quais são as grandes linhas de pesquisa. Como têm sido definidas e como se têm afirmado? Aqui, a liderança científica decerto tem uma palavra a dizer. Quais são as principais “necessidades estatísticas”? O conhecimento das respostas a essas questões informa como a comunidade dos estatísticos – a SPE – se vem afirmando.

A par dos textos produzidos pelos centros de investigação, no tema central deste boletim incluem-se artigos de opinião do actual e dos anteriores presidentes da SPE bem como de dois jovens investigadores com uma particularidade pois, de algum modo, estão “desbravando percursos inversos”. Tiago Marques, com formação de base em Biologia tem um doutoramento em Estatística (*Boletim SPE Primavera de 2008*, p. 49). Nuno Sepúlveda, Prémio SPE 2008 (divulgado no anterior Boletim), iniciado em Probabilidades e Estatística, está a terminar um doutoramento em Ciências Biológicas.

Foi enviado um convite a todos os centros onde se desenvolve investigação estatística. Imponderáveis de vária ordem não permitiram que alguns colaborassem neste número. Como responsável por esta edição SPE devo agradecer a generosidade de todos aqueles que, embora reconhecendo a importância desta iniciativa, não puderam aceitar o convite. Aqueles que, militantemente, disseram presente mostram o estado da arte neste domínio e porque primeiros, são exemplo que se regista; para mais tarde recordar. A eles, um agradecimento muito especial é devido. De facto, são centros de referência que historicamente também são génese; onde foram iniciados investigadores que já integram outros mais recentes e com novos projectos.

É o prosseguir da história com o contínuo avanço da Ciência Estatística, que se vai contando neste Boletim... já trintão!

O tema central do *Boletim Outono de 2009* será *Modelos Económicos*.



Mensagem do Presidente

Caros Colegas:

Com a regularidade a que o nosso colega Fernando Rosado nos tem habituado, aqui temos mais uma edição, desta vez primaveril, do nosso Boletim. Com ela, aqui vai também a tradicional Mensagem do Presidente, para vos anunciar as próximas actividades e para vos fazer um balanço breve do que se passou em 2008, mais em jeito de recordatória pois certamente têm recebido a informação que tem sido disponibilizada nos Boletins (incluindo este, isto no que se refere às notícias mais frescas), na página *web* ou por e-mail.

Após as eleições realizadas em Outubro, os membros eleitos da Mesa da Assembleia Geral, da Direcção e do Conselho Fiscal para o triénio 2009-11 tomaram posse a 22 de Janeiro de 2009, perante a Assembleia Geral. Estamos, pois, no início de um novo mandato, determinados a fazer todos os esforços para bem servir a SPE. Mas a construção da SPE é um trabalho colectivo dos seus sócios que continuará a requerer a vossa tradicional e empenhada participação.

A nossa colega Irene Oliveira capitaneou uma brilhante equipa da UTAD a quem confiámos a realização do nosso Congresso Anual, o XVI, realizado em Vila Real de 1 a 4 de Outubro de 2008. Publicaram-se, com o apoio do INE, as Actas do XV Congresso e o manual do minicurso do XVI Congresso (este sobre análise de dados espaciais, da autoria de Maria Lucília Carvalho e Isabel Natário). No âmbito da colaboração com o INE na publicação de segundas edições de manuais de minicursos já esgotados, saiu a primeira desta série, da autoria de Esmeralda Gonçalves e Nazaré Mendes Lopes, sobre séries temporais. Outras se seguirão brevemente.

Prosseguimos com os seminários dos novos doutores, continuámos os Encontros SPE-CIM (com uma pausa devida à mudança de Direcção do CIM), apoiámos várias reuniões científicas (com destaque para o COMPSTAT 2008, desta feita realizado no Porto e com a nossa colega Paula Brito a dirigir a equipa responsável pela sua excepcional organização), informámos sobre outras e sobre bolsas e oportunidades profissionais, colaborámos com várias entidades, mantivemos e melhorámos a página *web*, contribuímos para a melhoria de vários manuais escolares (com a revisão científica na área da Estatística), colaborámos no reajustamento dos Programas de Matemática do Ensino Básico. E demos prémios, o Prémio SPE 2008 (científico) e, com o apoio da Porto Editora, os Prémios Estatístico Júnior 2008 (para trabalhos de equipas de alunos do 3º ciclo do ensino básico e do ensino secundário), destinados à divulgação da Estatística junto da juventude. O nosso colega Dinis Pestana, que aceitou organizar para a SPE um Colóquio sobre Estatística e Sociedade, tema da maior relevância, brevemente vos dará notícias sobre a preparação desse evento. Outras iniciativas, no âmbito da Comissão Especializada de Educação (na área da divulgação da Estatística) e da Comissão Especializada de Nomenclatura Estatística (expansão do glossário estatístico), estão em marcha.

A Assembleia Geral da SPE de 22 de Janeiro deu, por grande maioria, luz verde a que nos associemos a outras sociedades estatísticas europeias para o lançamento de uma nova série de livros, “Studies in Theoretical and Applied Statistics”, da Springer, onde, se as negociações em curso chegarem a bom termo como tudo indica, passarão (entre outras publicações da iniciativa das sociedades participantes) a ser publicadas as Actas dos nossos Congressos. Para mantermos um registo em língua portuguesa suficientemente informativo da actividade dos nossos estatísticos, os resumos das comunicações dos Congressos, que são anualmente publicados, passarão a ter maior dimensão.

E, por falar em Congresso, é altura de os colegas iniciarem os preparativos para participarem no nosso XVII Congresso Anual, a realizar em Sesimbra de 30 de Setembro a 3 de Outubro de 2009, com os Colegas da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa a responsabilizar-se pela organização (ver informação neste Boletim). Ficaremos muito bem instalados, com excelentes vistas, e o programa é realmente aliciante. A vossa presença nesta celebração anual da Estatística em Portugal é indispensável. A SPE conta convosco.

Passado que foi 2008, é boa ocasião para nos lembrarmos de todas as pessoas e entidades que nesse ano apoiaram a SPE e as suas iniciativas e para lhes manifestar a nossa gratidão. Aos sócios quero agradecer o seu apoio e a sua participação na vida da Sociedade. Um agradecimento especial é devido àqueles sócios que aceitaram o pedido da Direcção para o exercício de diversas tarefas e funções.

Saudações cordiais



Notícias

• XVII Congresso SPE

O XVII Congresso Anual da Sociedade Portuguesa de Estatística, decorrerá no Sesimbra Hotel & SPA, de 30 de Setembro a 3 de Outubro de 2009, em Sesimbra. Co-organizado pelo Departamento de Matemática da FCT/UNL e pela Sociedade Portuguesa de Estatística e contará com a presença de cientistas de grande renome nacionais e estrangeiros. Assim, teremos o prazer de acolher como conferencistas plenários: Albert Shiryaev do Steklov Mathematical Institute, Moscovo, Rússia; Andrew Cairns da Heriot-Watt University Edinburgh, Reino Unido; Manuel Vilares do ISEGI/UNL; Paulo Eduardo Oliveira da FCT/UC e Tadeuz Calinski do Department of Statistical Methods, Agricultural University of Poznan, Polónia. O mini-curso sobre Análise de sobrevivência será da responsabilidade de Cristina Rocha da FC/UL. É pois com a maior satisfação que convidamos a comunidade Estatística a participar neste evento na vertente científica, sob a forma de comunicações orais ou posters. Informação mais detalhada, sobre esta edição do Congresso, poderá ser obtida brevemente na página <http://www.fct.unl.pt/spe2009>.

Contamos com a vossa presença para face ao maravilhoso Atlântico aproveitarmos mais esta oportunidade de qualificação, intercâmbio e convívio científicos. Sublinhamos que a data limite para a apresentação dos resumos é 15 de Maio de 2009.

Pela Comissão Organizadora Local,
Manuel L. Esquível e Isabel Natário

• Estatística em Educação

No passado mês de Novembro foi lançada a publicação *Measuring Improvements in Learning Outcomes* (http://www.oecd.org/document/54/0,3343,en_2649_39263231_41701046_1_1_1_37455,00.html) pela Organização para a Cooperação e Desenvolvimento Económico (OCDE). A definição de modelo de valor acrescentado adoptada no relatório é a seguinte:

“A class of statistical models that estimate the contributions of schools to student progress in stated or prescribed education objectives (e.g.cognitive achievement) measured at at least two points in time.”

Governos de diversos países têm manifestado interesse crescente em implementar os modelos de valor acrescentado com vista a obter indicadores mais rigorosos da contribuição das escolas e/ou professores nos resultados escolares atingidos pelos alunos. O principal objectivo do relatório é contribuir para difundir o estado da arte desta classe de modelos, discutir potencialidades e limitações e apresentar algumas práticas.

O relatório estrutura-se em três partes orientadas a públicos diferenciados. Na primeira discutem-se os objectivos e o uso da modelação de valor acrescentado, incluindo aqui os objectivos numa perspectiva de política educativa, a natureza dos resultados que devem ser de domínio público e formas de divulgação, bem como a sua utilização para promover e melhorar a educação. A segunda parte discute aspectos metodológicos e técnicos dos modelos, ao mesmo tempo que identifica quais são os requisitos chave que necessitam de ser considerados. Também são apresentados alguns exemplos. A terceira parte está centrada nas fases da implementação deste tipo de modelos e constitui uma ajuda aos governantes e dirigentes que efectivamente pretendam implementá-los nos seus países.

Maria Eugénia Ferrão

• Curso Satélite - Análise de Sobrevivência em Estudo de Coorte

Realizou-se, em Novembro passado, na VI Conferência de Estatística e Qualidade na Saúde (EQS2008). Este curso, de carácter interdisciplinar, foi ministrado por Valeska Andreozzi, investigadora auxiliar no CEAUL, no âmbito do Ciência 2007 com base no livro “*Análise de Sobrevivência: Teoria e Aplicações em Saúde*”, 2005, Editora Fiocruz.

O objectivo principal do curso foi o de propiciar ao público alvo (alunos e investigadores da área da saúde e estatística) os conceitos teóricos e práticos da análise de sobrevivência que pudessem ser aplicados na sua própria investigação. Foram apresentadas, de forma simples e acessível, as várias formas de modelar o tempo de sobrevivência na área da saúde. Foram abordados os modelos de Cox e as suas extensões para tratar covariáveis dependentes do tempo, eventos múltiplos e fragilidade (efeitos aleatórios). Ilustrações foram feitas através de dados reais de estudos de coorte e do software R.

O curso contou com a presença de alunos da área da saúde e estatística, bem como professores e investigadores de ambas as áreas científicas e teve lotação ao redor de 50 participantes.

Os slides do curso estão disponíveis sob requisição via e-mail: valeska.andreozzi@fc.ul.pt. Para maiores informações e material didático ver página do livro <http://dengue.procc.fiocruz.br/~sobrevida/>

Valeska Andreozzi

• 2º Fórum do ALEA

Realizou-se no dia 30 de Janeiro, na Escola Vasco da Gama, em Lisboa, o 2º Fórum do ALEA, que teve como tema *Numeracia e Cidadania*. Foi uma reunião muito concorrida, em que estiveram presentes mais de 200 professores, do Ensino Básico e Secundário.



Na divulgação da iniciativa lia-se: “Este fórum visa agregar docentes, investigadores, alunos, técnicos, etc., de todas as áreas do conhecimento, sendo autores e/ou utilizadores do ALEA, ou que de algum modo contactem com a Estatística”.

Os tópicos abordados

- Literacia Estatística;
- Estatística, Estatísticas e Tecnologias de Informação e Comunicação;
- Informação para a cidadania;
- Ensino e Aprendizagem da Estatística e Probabilidades;
- Experiências e Projectos Interdisciplinares com o ALEA.

foram desenvolvidos em dois painéis, pelos seguintes oradores:

1º painel: Ter Estatística nos currícula faz os cidadãos terem mais literacia?

Eugénia Graça Martins (FCUL) - *O tema de Estatística e Probabilidade no (Reajustamento do) Programa de Matemática do Ensino Básico*

Luísa Castro Loura (SPE) - *A Estatística e os grandes temas da actualidade*

Maria João Lagarto (GAVE-ME) - *O que fazer com estes dados?*

2º painel: Ensinar e aprender Estatística no Ensino Básico e Secundário

Isabel Catalão (EBI Quinta de Marrocos) - *Percursos com o ALEA*

José Paulo Viana (ES Virgílio Ferreira) - *Problemas e Estatística*

Emília Oliveira/Pedro Campos (Projecto ALEA) - *Ensinar e aprender Estatística com o ALEA.*

Na sessão de abertura estiveram presentes a Directora Regional de Educação do Norte, Margarida Moreira, assim como a Presidente do INE, Alda Carvalho, tendo a conferência de abertura sido proferida pelo director do Plano Tecnológico para a Educação, João Mata.

A sessão decorreu num ambiente em que se denotou o interesse por parte dos professores, por iniciativas deste género.

Eugénia Graça Martins e Luísa Loura

• Alteração do modelo de Actas dos Congressos Anuais da SPE

No passado dia 22 de Janeiro, por solicitação da Direcção, reuniu a Assembleia Geral da Sociedade Portuguesa de Estatística. O ponto fundamental da agenda foi a alteração do modelo de actas dos congressos anuais. Para esta importante reunião os sócios acorreram em número revelador da importância do tema em debate.

Introduzindo a temática o Presidente da SPE tinha enviado aos sócios o seguinte Memorando:

“As Actas dos Congressos Anuais da SPE seguem um modelo estabelecido há muitos anos, que certamente conhecem bem, e que foi confirmado na Assembleia Geral realizada em 2005 em Tomar. Recebemos muito recentemente uma proposta da Sociedade Italiana de Estatística relativo a uma nova “international book series in Statistics” da Springer, cujo título provisório é “Studies in Theoretical and Applied Statistics”, na qual, além de outras possíveis publicações de interesse para as Sociedades Estatísticas fundadoras, seriam publicadas as Actas dos seus Congressos. Há aspectos ainda a negociar, mas exporei as principais ideias propostas. Os “Managing Editors” da série seriam em princípio os Presidentes das sociedades participantes e o “Editorial Board” seria constituído por nomes relevantes propostos pelas sociedades fundadoras. Cada volume de Actas teria os seus editores próprios indicados pela respectiva sociedade mais um editor da série não pertencente à sociedade organizadora do Congresso. Torna-se agora possível incluir avaliadores de outros países nos painéis de avaliadores das Actas. As Actas passariam a ser em inglês com um limite máximo de 8 a 12 páginas por artigo e cerca de 400 a 500 páginas. Teríamos de assegurar a compra de 500 exemplares (aspecto a negociar porque nós só precisaríamos de cerca de 350 para distribuir pelos participantes e para oferecer a bibliotecas). O custo previsto por exemplar é, no caso de 400 páginas, de 19 Euros. As Sociedades italiana e francesa já manifestaram a sua vontade de aderir e é-nos perguntado se queremos também estar entre as Sociedades fundadoras, decisão que tem de ser tomada sem demora. Esta seria uma de várias formas de cooperação entre sociedades estatísticas europeias que se poderiam vir a promover.

As principais mudanças nas Actas referem-se à língua, que passaria a ser o inglês, e ao facto de os artigos nelas publicados passarem a ser divulgados internacionalmente. Há também algumas pequenas diferenças na edição e avaliação, que passariam a ter alguma internacionalização, e nos custos, que seriam superiores mas perfeitamente comportáveis no âmbito dos orçamentos dos Congressos. Para se manter um registo em português da actividade científica nacional, pensamos que seria conveniente os resumos das comunicações aos congressos (que são distribuídos aos congressistas no início do congresso no “Programa e Livro de Resumos”) serem mais alargados, passando a ter duas páginas e continuando naturalmente a ser em português.

A Direcção, na sua última reunião de 24 de Novembro apreciou favoravelmente a proposta, mas entende que, pela importância do assunto para a vida futura da SPE, deve ser a Assembleia Geral a decidir, tanto mais que o actual modelo das Actas dos Congressos foi confirmado em Assembleia Geral em 2005 e não foi proposta qualquer modificação no programa de recandidatura da Direcção recentemente reeleita.

A Mesa da Assembleia Geral decidiu-se pela convocação da Assembleia para uma reunião cujo principal objectivo é o de discutir esta matéria e de sobre ela deliberar.

Dado que se trata de matéria extremamente importante para o futuro da SPE, apelamos a todos os sócios para fazerem um especial esforço e participarem nesta reunião.

Cordiais cumprimentos / Carlos Braumann / Presidente da SPE.”

Debatida a questão em análise, a Assembleia, por larga maioria, votou favoravelmente a alteração proposta.

A Direcção implementará a decisão tomada e o Presidente da SPE negociará, com as demais sociedades parceiras e com a editora, no sentido de adaptar essa nova edição ao caso português.

Fernando Rosado

• Probabilidades e Estatística no Mat - Oeste

No dia 17 de Julho de 2009 irá decorrer, na Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Leiria, a segunda edição do Mat-Oeste - Matemática na Região Oeste (informação detalhada sobre este encontro pode ser consultada no *site* www.estg.ipleiria.pt/~mat-oeste). O evento é dirigido a docentes de todos os níveis de ensino e ao público em geral, pretendendo ser um encontro onde se discutam e troquem ideias e experiências nas mais diferentes vertentes da Matemática, procurando promovê-la na Região Oeste.

Em cada ano o Mat-Oeste é subordinado a uma área específica da Matemática, sendo esta segunda edição dedicada ao tema das Probabilidades e Estatística. Por tal razão, neste evento contaremos com a presença de diversas personalidades de destaque na divulgação, no ensino e na investigação em Probabilidades e Estatística, estando já confirmada a presença do Prof. Doutor Nuno Crato.

Para além das sessões de comunicações orais, os participantes terão ainda a possibilidade de assistir à entrega dos prémios da 1.ª edição do Prémio Doutor Pedro Matos (sobre o qual é possível colher informações mais precisas no *site* www.premiopedromatos.ipleiria.pt). Este certame, destinado a alunos do secundário, terá todos os anos um tema distinto, derivado duma área específica da Matemática, tendo sido nesta primeira edição eleita a área das Probabilidades e Estatística. O objectivo principal deste prémio é não só fomentar a criatividade e o interesse pela Matemática e as suas aplicações, bem como contribuir para que despontem novos jovens talentos. Neste sentido, podem candidatar-se ao Prémio estudantes do ensino secundário, individualmente ou em grupo com o máximo de três estudantes. Do grupo pode ainda fazer parte um professor do ensino secundário, ao qual caberá o papel de orientador. Além da entrega dos prémios, durante o encontro estarão expostos os Posters associados aos melhores trabalhos candidatos ao Prémio.

Rui Santos

• Revista TMQ - Qualidade

Vai ser publicado, através da Sílabo, no primeiro trimestre deste ano o número zero de uma nova revista denominada TMQ Qualidade no âmbito da produção editorial do Departamento de Métodos Quantitativos da ISCTE Business School, especificamente, na denominada série Temas em Métodos Quantitativos (TMQ). Esta revista, prosseguindo a adoptada perspectiva de multidisciplinaridade, pretende aprofundar as várias vertentes da temática da Qualidade e, assim, contribuir para uma mais completa informação sobre conhecimentos, práticas e opiniões de profissionais e académicos de várias áreas, constituindo um suporte e incentivo à aplicação, investigação e elaboração de dissertações de mestrado e teses de doutoramento nestes domínios.

Empresários, gestores, docentes, alunos e todos os interessados na Qualidade encontrarão nesta obra um fórum de exposição das suas ideias e de consulta sobre esta temática, onde se incluem experiências pessoais e visões concordantes ou críticas, genéricas ou específicas, voltadas para a filosofia e as ferramentas da Qualidade, podendo os autores optar por uma submissão sujeita ou não a referees.

Está em fase de desenvolvimento um portal da revista. Quem estiver interessado em receber informações sobre este projecto poderá solicitá-las enviando uma mensagem para os endereços msaraiva@uevora.pt ou afst@iscte.pt.

Os próximos números terão duas partes, sendo a primeira dedicada a um tema específico (no primeiro número a Qualidade no Sector da Saúde), e a segunda a todos os temas, incluindo os métodos quantitativos.

Está já aberta a fase de “call for papers” para o primeiro número regular, com a data limite de 15 de Setembro, devendo os interessados contactar os editores através dos endereços acima apresentados.

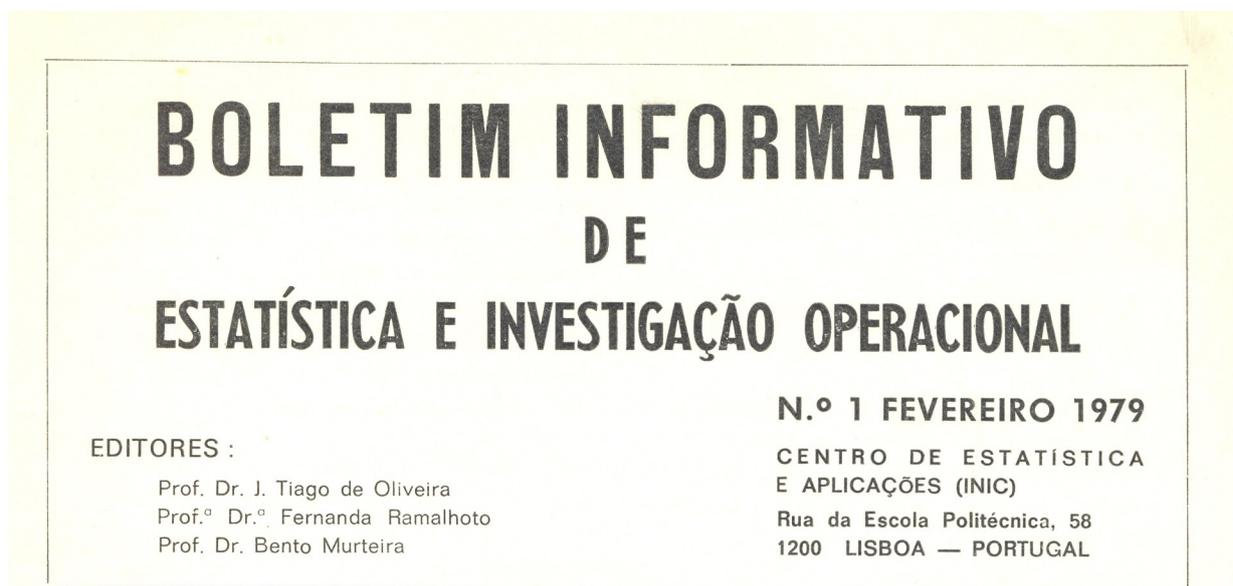


António Teixeira

Episódios na História da Estatística

• 30 anos de Boletim

O Boletim Informativo de Estatística e Investigação Operacional já criou um anterior episódio (Boletim Primavera de 2007, p. 14). Desta vez, aniversariante, é revisitado.



O Boletim Primavera de 2009 vai entrar no prelo em Fevereiro; mês em que, há 30 anos, nasceu o seu predecessor Boletim Informativo de Estatística e Investigação Operacional.

É o momento para, reconhecidamente, homenagear os seus criadores e dinamizadores.

Sabe-se da extrema importância dessa iniciativa que simboliza o despertar de Portugal para uma nova era científica, também na área da Estatística.

Porquê um Boletim? O Editorial número 1 responde sucintamente e fornece uma lista de objectivos que, na generalidade, podemos considerar actuais.

O Boletim número 1 apresenta também um artigo de fundo da autoria de J. Tiago de Oliveira – o primeiro Presidente da SPE e o principal dinamizador na criação da Sociedade Portuguesa de Estatística. Esse texto intitulado “Alguns núcleos recentes de investigação estatística em Portugal” é uma boa síntese nas origens da SPE e, por consequência, uma referência na história da Estatística e da Investigação em Portugal.

Como era e como é? Para comparar (e reflectir!) é relevante a página 5 do Boletim 1 - abaixo, em edição *fac-simile* - onde se salientam “algumas das principais revistas de Estatística e Investigação Operacional existentes no Instituto Gulbenkian de Ciência – Biblioteca do Centro de Cálculo Científico e Biblioteca do Centro de Estudos de Economia Agrária”. Este, que agora seria um pequeno detalhe de informação, era considerado à época, com direito a publicação, para conhecimento dos leitores. Na época relatada, de facto, o Centro de Cálculo da Gulbenkian era “o único” para apoio aos investigadores – desde a biblioteca ao computador. Os nossos resultados computacionais actualmente estão “à distância de um clique”. Na realidade, há trinta anos, o centro de cálculo da Gulbenkian ainda era “o pólo informático”, quase único, onde “alguns”, entre os quais os alunos finalistas da FCUL, podiam entregar os seus programas computacionais que eram executados pelos operadores do centro e sobre os quais recebiam os respectivos “outputs no dia seguinte” – um extenso volume de papel, muitas vezes apenas com o registo da não execução devido a erro de programação. A palavra package (e muito menos o pacote estatístico!) ligada à ferramenta de tão fácil acesso que actualmente usamos ainda não tinha nascido. Tudo está diferente...

Fernando Rosado

BOLETIM INFORMATIVO DE ESTATÍSTICA E INVESTIGAÇÃO OPERACIONAL

EDITORES :

Prof. Dr. J. Tiago de Oliveira
Prof.ª Dr.ª Fernanda Ramalhoto
Prof. Dr. Bento Murteira

N.º 1 FEVEREIRO 1979
CENTRO DE ESTATÍSTICA
E APLICAÇÕES (INIC)
Rua da Escola Politécnica, 58
1200 LISBOA — PORTUGAL

EDITORIAL

Pareceu-nos ser importante para uma melhor aproximação entre as pessoas de qualquer modo ligadas à Estatística e Investigação Operacional, a publicação periódica de um "Boletim Informativo" sobre o que se vai passando no país e no estrangeiro nestes domínios.

Estamos convictos que neste momento há já no país um grupo considerável de investigadores, docentes e diferentes utilizadores de Estatística e Investigação Operacional que sentem a necessidade de uma melhor utilização dos nossos recursos humanos e técnicos, embora tais recursos a nível mundial ou mesmo europeu sejam ainda surpreendentemente modestos. Pensamos também que este Boletim pode vir a servir de trampolim para uma futura colaboração científico-tecnológica entre Portugal, Brasil e outros países. Tal intercâmbio poderá fomentar a formação de "grupos de trabalho" em Estatística e Investigação Operacional quer a nível da investigação pura quer no campo da aplicação a problemas concretos de técnicas científicas.

No fundo através do Boletim queremos fornecer os meios para :

— Pôr questões, ou responder a questões postas por outros leitores do Boletim, sobre problemas de Estatística ou Investigação Operacional de índole teórica ou prática.

— Fazer relatos sobre Conferências nacionais ou estrangeiras em que tenha participado e/ou colaborado.

— Fazer um apelo para uma possível ajuda num determinado trabalho científico ou prático, indicando condições.

— Anunciar Conferências, Congressos, Encontros, Seminários, Cursos na Universidade ou fora dela quer no país quer no estrangeiro.

— Iniciar discussão entre os leitores do Boletim sobre o ensino da Estatística e Investigação Operacional no Ensino Secundário.

— Anunciar empregos no campo de Estatística Investigação Operacional e Computação na Universidade e fora dela.

— Anunciar livros e revistas científicas ou de divulgação de Matemática, Estatística, Investigação Operacional, Computação, a publicar brevemente ou recentemente adquiridos por uma determinada biblioteca.

— Publicar crítica a livro ou revista científica ou de divulgação, desde que não sejam excedidas as 1.000 palavras.

— Sugerir rúbricas que não estejam aqui contempladas.

Enfim estamos certos da necessidade deste Boletim, acreditamos vivamente que a divulgação da informação é algo de muito importante.

O Boletim poderá ser algo de negativo ? !
Alea jacta est !

FRANCA RAMALHOTO

Todas as contribuições ou pedidos de informação a este Boletim devem ser enviadas a :

Fernanda Ramalhoto

Centro de Estatística e Aplicações da Universidade de Lisboa (INIC)

R. da Escola Politécnica, 58
1200 Lisboa — Portugal

As contribuições devem ser dactilografadas e não devem nunca exceder as 1.000 palavras. Sempre que possível agradecemos, por questões óbvias de espaço, que não fossem excedidas as 250 palavras.

Quanto à periodicidade deste Boletim, será em princípio quadrimestral, podendo tornar-se trimestral se o acolhimento e contribuições o justificarem. Todas as notícias, artigos, cartas etc., que se deseja venham a ser publicadas num dado número deste Boletim, devem chegar à morada acima indicada pelo menos um mês antes da sua data de publicação.

"ALGUNS NÚCLEOS RECENTES DE INVESTIGAÇÃO ESTATÍSTICA EM PORTUGAL"

Talvez tenha algum interesse referir, em curso rápido, a pesquisa, com componentes teóricas, que recentemente se tem feito entre nós, após a II Guerra Mundial.

Essencialmente parece que os núcleos fundamentais de estudos têm estado ligados ao Instituto Superior de Ciências Económicas e Financeiras, à Faculdade de Ciências de Lisboa e, também, de índole bastante mais aplicada, ao Instituto Nacional de Estatística. Referimos, para breve notícia, que há alguns estudos individuais de investigação aplicada e ainda de outras organizações, como o Banco de Portugal e também grupos incipientes.

A actividade do Centro de Estudos de Estatística Económica (do IAC) ligado a gente do ISCEF, e que se estende dos anos 50 a meados dos anos

60, foi essencialmente voltada ao estudo de metodologia ligada à pesquisa económica (em grande parte crono-séries e programação) e um pouco à pesquisa demográfica. Este núcleo, durante a sua experiência, à volta de uma dúzia de anos, trouxe a Portugal alguns investigadores de renome. Deve ainda salientar-se o estudo feito por 1963 pelo Centro relativo às necessidades educacionais para 1975, integrado no Projecto Regional do Mediterrâneo.

Virado mais à investigação teórica, desde meados dos anos 50, no Seminário de Matemática, depois no Centro de Matemáticas Aplicadas e, actualmente, no Centro de Estatística e Aplicações (do IAC e, agora do INIC) a gente da Faculdade de Ciências de Lisboa e outros têm prosseguido estudos, essencialmente em teoria das probabilidades, estatística de extremos, processos estocásticos e planeamento de experiências, além do apoio consultivo a investigadores em Biologia, Medicina, Psicologia, História, etc. O trazer de alguns investigadores e a realização de cursos têm sido actividades, também, dos sucessivos centros. Actualmente, no CEA, as linhas de pesquisa são a decisão estatística, em particular relativa a extremos estatísticos, processos estocásticos e aplicações, planeamento de experiências e programação e aplicações.

Constituídos em meados dos anos 40 os Centros de Estudos Económicos e Demográficos do Instituto Nacional de Estatística têm executado alguns estudos ligados à sua definição. Deverão recordar-se, entre outros, os estudos de demografia portuguesa realizados.

É bem pouco, quer na área de investigação teórica, quer na das aplicações o que entre nós se tem realizado, muitas vezes fruto do interesse e persistência individual. Países da dimensão do nosso têm, há muito, Institutos de Estatística nas Universidades em ligação com os Institutos de Aplicações. Estamos bem atrasados ! Temos, pois, de avançar rapidamente para o futuro, para o que temos gente capaz. Não se pode perder mais tempo !

J. TIAGO DE OLIVEIRA

2

ESTRANGEIRO

COMPUTAÇÃO ESTATÍSTICA

Com o sentido de informar e de, ao mesmo tempo, tentar mostrar a necessidade de colaboração a nível país, passamos à seguinte notícia vinda do estrangeiro.

Durante a 41.ª Sessão do ISI (Instituto Internacional de Estatística) a Assembleia Geral deste Instituto aprovou a formação de uma Associação Internacional de Computação Estatística (IASC, International Association for Statistical Computing). Os objectivos desta Associação são os de fomentar à escala mundial o interesse efectivo pela Computação Estatística e facilitar a permuta dos conhecimentos técnicos já adquiridos, através de contactos internacionais e encontros entre estatísticos, entre profissionais de computação, entre organizações, entre instituições e até mesmo entre governos.

No sentido de reduzir ao mínimo a burocracia no seio da recém formada Associação foi adoptado o sistema de formação de "grupos de trabalho" em vez de Comissões permanentes. Esses "grupos de trabalho" são estabelecidos por um curto espaço de tempo que não pode ultrapassar os dois anos. Qualquer membro desta Associação pode pedir o estabelecimento de um "grupo de trabalho" submetendo para isso, ao presidente da Associação, um projecto escrito, descrevendo pormenorizadamente os objectivos do grupo, para que tal projecto possa ser convenientemente examinado pela Comissão Executiva da Associação. Desde que esta Comissão dê o seu aval a tal projecto o "grupo de trabalho" requerido será formado. Por outro lado o Boletim Informativo ISI compromete-se a publicar um anúncio nas suas colunas, convidando todos os membros da Associação interessados nesse tipo de projecto a contactarem o chefe do grupo, afim de colher mais informações e eventualmente vir a fazer parte desse tal grupo recém formado.

Para que todos os membros da Associação possam ter uma ideia clara do trabalho desenvolvido pelos diversos grupos, cada "grupo de

trabalho" tem de elaborar um relatório anual detalhado sobre o trabalho desenvolvido durante esse ano.

Neste momento foram já organizados 13 grupos, com as actividades seguintes:

- Data Editing/Imputation
- Evaluations (Software and Hardware)
- Hand-held Programmable Calculators
- Input/Output Activities
- Languages
- Policy and Procedures for Acquisition and Management of Computing
- Portability
- Privacy and Confidentiality
- Security
- Simulation
- Statistical Data Bases and Data Base Systems
- Survey and Census Processing
- Unique Requirements of Developing Countries

(Usámos a nomenclatura inglesa por não estarmos certos das respectivas traduções em Português).

Fazem parte desta Associação - entre outras, entidades dos Estados Unidos, Holanda, Nova Zelândia, Austrália, Índia, Chile, Checoslováquia, Japão, Alemanha Federal e Democrática e Inglaterra.

FRANCA RAMALHOTO

COMENTARIO E SUGESTÃO

Aproveitamos a deixa do artigo anterior para lembrar a necessidade, em termos de independência linguística, de encargar a sério a tarefa árdua mas imperiosa de criar um vocabulário de termos tecnológicos em Português.

3

CALENDÁRIO INTERNACIONAL

— A "3rd International Conference on Application of Statistics and Probability in Soil and Structural Engineering" realizar-se-á de 29 de Janeiro a 2 de Fevereiro de 1979 em Sidney, Austrália.

— "Encontro Internacional de Crono-Séries", a realizar de 27 a 29 de Março de 1979, em Nottingham, Inglaterra.

— A "2.ª Conferência sobre Fiabilidade" realizar-se-á de 28 a 30 de Março de 1979, no Hotel Metrôpole, no "National Exhibition Centre", Birmingham, Inglaterra.

— A "Royal Statistical Society" leva a cabo uma conferência de 2 a 5 de Abril de 1979, na Universidade de Oxford, Inglaterra. Resumo dos principais assuntos a serem tratados nesta conferência: Exploratory Data Analysis; Time Series; Scaling and Robustness; Statistics for Managers; Social Statistics; Compound Distributions; Contingency Tables. Colaboram entre outros os professores: M. Stone, P. Armitage, A. F. M. Smith, D. J. Bartholomew, D. F. Hendry e G. E. P. Box.

— Uma breve conferência sobre Genstat (Programas ligados à análise estatística) realizar-se-á de 9 a 11 de Abril de 1979, no "Computer Laboratory Cambridge", Inglaterra.

— A UNESCO realiza em Paris no mês de Abril o "Seventh Meeting of the Panel Experts to discuss the draft of the Manual of Application of Standardized Concepts in Collection of Science and Technology Statistics". Para mais informações sobre este encontro, não público, escrever para:

Director, Office of Statistics, UNESCO
7, Place de Fontenay
75700 Paris
FRANÇA

— A "Associação das Sociedades Europeias de Investigação Operacional" sob os auspícios da IFORS, leva a cabo o "3.º Congresso Europeu de Investigação Operacional" de 9

a 11 de Abril de 1979, em Amsterdão, Holanda.

— "VI Jornadas Luso-Espanholas de Matemática", 1979. Está prevista a sua reabilitação em Santander, Espanha, para depois da Páscoa.

— A UNESCO realiza em Maio de 1979 em Viena, Austria, um encontro intitulado "Working Group on Statistics and Indicators of Culture". Para detalhes sobre este encontro, também não público, contactar a morada do anúncio referente ao encontro da UNESCO a realizar em Paris em Abril de 1979.

— O "Computer Science and Statistics, 12th Annual Symposium on the Interface" realizar-se-á de 10 a 11 de Maio de 1979, na Universidade de Waterloo, Waterloo, Ontário, Canadá.

— A "4.ª Escola Internacional de Verão sobre Problemas de Escolha de Modelos e Estimação de Parâmetros em Análise de Regressão" será levada a cabo de 7 a 16 de Maio de 1979, em Berlim na República Democrática Alemã.

— A "9.ª Conferência sobre Processos Estocásticos e Aplicações", patrocinada pela "Bernoulli Society for Mathematical Statistics and Probability", realizar-se-á de 6 a 10 de Agosto de 1979, em Evanston, Illinois, Estados Unidos da América.

— O "12.º Encontro Europeu de Estatísticos", patrocinado pela "Bernoulli Society for Mathematical Statistics and Probability" realizar-se-á de 4 a 7 de Setembro de 1979, em Varna, Bulgária.

— A "23.ª Conferência Anual da Organização Europeia para o Controlo da Qualidade" realizar-se-á de 4 a 7 de Setembro de 1979, em Budapeste, Hungria.

— O "Encontro Europeu da Sociedade de Econometria" realizar-se-á de 4 a 7 de Setembro de 1979, em Atenas na Grécia.

— A "42.ª Sessão Bial do Instituto Internacional de Estatística" realizar-se-á de 4 a 14 de Dezembro de 1979, em Manila nas Filipinas.

4

ACORDOS CULTURAIS (INIC)

BRASIL

Oferta anual e recíproca de 5 bolsas, no máximo com 45 mensalidades e por um período mínimo de 3 meses. Boas perspectivas a nível da investigação científica e técnica nos campos da computação, biologia, medicina, engenharia, pecuária, vitivinicultura, melhoria de sementes e administração pública e privada.

BULGARIA

Apoio à cooperação científica entre as instituições e centros de investigação, no âmbito do ensino superior, através de intercâmbio de docentes e investigadores; organização de estágios, seminários, conferências, troca de publicações e de documentação e programas de investigação em comum.

CHECOSLOVAQUIA

Concessão anual de bolsas de Estudo que vão de 10 dias a 1 ano académico, para participação em congressos e missões de estudo a nível universitário, incluindo frequência de cursos nas Universidades checoslovacas.

ESPAÑHA

Interesse em concretizar o intercâmbio nos domínios da investigação científica. Concessão de bolsas de estudo destinadas a estudos de pós-graduação e trabalhos de investigação. As condições modalidades e quantias das bolsas serão estabelecidas por via diplomática.

FRANÇA

Interesse em intensificar os contactos entre organismos especializados e universitários dos dois países. Organização de colóquios bilaterais e de programas de trabalho sobre temas definidos de comum acordo. Concessão de bolsas de estudo de 8 meses de duração a investigadores portugueses e franceses.

Existem ainda acordos culturais entre Portugal (INIC) e Grã-Bretanha e Irlanda do Norte, Holanda, Hungria, Jugoslávia, Polónia, Senegal e União Soviética, sobre os quais nos debruçaremos no próximo número deste Boletim. Todos os assuntos referentes a estes acordos são tratados na Divisão de Planeamento e Intercâmbio do INIC.

REVISTAS

São as seguintes, algumas das principais revistas de Estatística e Investigação Operacional existentes no INSTITUTO GULBENKIAN DE CIENCIA em Oeiras.

BIBLIOTECA DO CENTRO DE CALCULO CIENTIFICO

CANADIAN MATHEMATICAL BULLETIN
1964/7
1968/11 a 1974/17

COLLOQUIUM OTAN RECHERCHE OPERATIONELLE (BULL. D'INFORMATION)
1964/2 e 3

INTERNATIONAL ABSTRACTS IN OPERATIONS RESEARCH (BALTIMORE)
1974/14 a 1978/18

JOURNAL OF MATH. AND PHYSICS
1963/42 a 1968/47

(continuado por: Studies in Applied Mathematics)

JOURNAL OF OPERATIONAL RESEARCH SOCIETY (OXFORD)
1978/29

(continuado de: Operational Research Quarterly)

JOURNAL OF THE SIAM
1953/1 a 1965/13

(continuado por: Siam Journal on Applied Mathematics)

JOURNAL OF THE SIAM SERIES A: CONTROL
1962/1 a 1965/3

(continuado por: Siam Journal on Control)

JOURNAL OF THE SIAM SERIES B: NUMERICAL ANALYSIS
1964/1 a 1965/3

(continuado por: Siam Journal on Numerical Analysis)

JOURNAL OF STATISTICAL COMPUTATION AND SIMULATION
1976/5 a 1978/8

MATHEMATICS OF COMPUTATION (PROVIDENCE R. I.)
1960/14 a 1978/32

(continuação de: Mathematical Tables and other Aids to Computation)

MATHEMATICAL PROCEEDINGS OF THE CAMBRIDGE PHILOSOPHICAL SOCIETY (CAMBRIDGE)
1975/77 a 1978/84

(continuação de: Proceedings of the Cambridge Philosophical Society)

MATHEMATICAL PROGRAMMING (AMSTERDÃO)
1971/1
1973/4 a 1978/15

MATHEMATICAL PROGRAMMING STUDY (AMSTERDÃO)
1975/4 a 1978/9

MATHEMATICAL REVIEWS (ANN. ARBOUR MICH.)
1968/35 a 1978/56

MATHEMATICAL SYSTEMS THEORY (N. YORK)
1967/1 a 1976/10

MATHEMATICAL TABLES AND OTHER AIDS TO COMPUTATION
1943/1 a 1959/13

(continuado por: Mathematics of Computation)

METRA (PARIS)
1964/3 a 1975/14

MICHIGAN MATHEMATICAL JOURNAL (ANN. ARBOUR MICH.)
1963/10 a 1978/25

OPERATIONAL RESEARCH QUARTERLY (OXFORD)
1965/16
1973/24 a 1978/28

(continuado por: The Journal of the Operational Research Society)

OPERATIONS RESEARCH (BALTIMORE)
1974/22 a 1978/26
(FALTA 1977)

PROCEEDINGS OF THE CAMBRIDGE PHILOSOPHICAL SOCIETY
1954/50
1963/59 a 1974/76

(continuado por: Mathematical Proceedings of the Cambridge Philosophical Society)

PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON OPERATIONS RESEARCH
1963 - 3.º
1966 - 4.º
1969 - 5.º
1972 - 6.º
1975 - 7.º

PROCEEDINGS OF THE SYMPOSIUM ON CONGESTION THEORY
1964

PUBLICATIONS DE L'INSTITUT DE MATHEMATIQUE
1967/7 a 1975/18

PUBLICATIONS DE L'INSTITUT DE STATISTIQUE DE L'UNIVERSITÉ DE PARIS
1963/12 a 1971/20

PUBLICATIONS MATHEMATIQUES (PARIS)
1964/23
1965/26
1967/33

5

6

QUARTERLY OF APPLIED MATHEMATICS (PROVIDENCE R. I.)
1963/21 a 1978/36

RAIRO — Analyse Numerique/Numerical Analysis (AF CET)
1977/11 a 1978/12

RAIRO — Automatique/Systems Analysis and Control (AF CET)
1977/11 a 1978/12

RAIRO — Informatique/Computer Science (AF CET)
1977/11 a 1978/12

RAIRO — Recherche Operationelle/Operations Research
1977/11 a 1978/12

REVUE FRANÇAISE D'INFORMATIQUE ET RECHERCHE OPERATIONELLE
Serie Bleue 1967/1 a 1976/10
Serie Verte 1968/2 a 1976/10
Serie Rouge 1967/1 a 1976/10
Serie R (Informatique) 1974/8 a 1976/10
(continuada pelas: RAIRO)

REVUE FRANÇAISE DE RECHERCHE OPERATIONELLE
1963/7 a 1966/10

REVUE DE MATHEMATIQUES PURES ET APPLIQUEES (BUCAREST)
1963/8

(continuado por: Revue Roumaine de Mathematiques Pures et Appliquees)

REVUE ROUMAINE DE MATHEMATIQUES PURES ET APPLIQUEES (BUCAREST)
1964/9 a 1978/23
(continuado de: Revue de Mathematiques Pures et Appliquees)

SIAM JOURNAL ON COMPUTING (PHILADELPHIA)
1972/1 a 1978/7

SIAM JOURNAL ON APPLIED MATHEMATICS (PHILADELPHIA)
1966/14 a 1978/35

(continuado de: Journal of the SIAM)

SIAM JOURNAL ON CONTROL (PHILADELPHIA)
1966/4 a 1975/13

(continuado de: Journal of the SIAM, Series A: Control)
(continuado por: SIAM Journal on Control and Optimization)

SIAM JOURNAL ON CONTROL AND OPTIMIZATION
1976/14 a 1978/16

(continuado de: SIAM Journal on Control)

SIAM JOURNAL ON MATHEMATICAL ANALYSIS (PHILADELPHIA)
1970/1 a 1978/9

SIAM JOURNAL ON NUMERICAL ANALYSIS (PHILADELPHIA)
1966/3 a 1978/15

(continuado de: Journal of the SIAM Series B: Numerical Analysis)

SIAM REVIEW
1965/7 a 1978/20

SIAM NEWS
1976/9 a 1978/11

SCIENCE CITATION INDEX DESDE 1971

SOVIET MATHEMATICS (PROVIDENCE R. I.)
1965/6 a 1978/19

STOCHASTIC PROCESSES AND THEIR APPLICATIONS (AMSTERDÃO)
1973/1 a 1978/8

STUDIA SCIENTIARUM MATHEMATICARUM HUNGARICA
1966/1 a 1974/9

STUDIES IN APPLIED MATHEMATICS
1969/48 a 1972/51
(continuado de: Journal of Math and Physics)

THEORY OF PROBABILITY AND ITS APPLICATIONS (SIAM; RUSSIAN JOURNAL; PHILADELPHIA)
1966/11 a 1977/22

TRABAJO DE ESTADISTICA Y INVESTIGACION OPERATIVA (MADRID)
1962/13 a 1977/28

UNIVERSITY OF CALIFORNIA PUBLICATIONS IN STATISTICS
1954 a 1956/2
1960 a 1963/3

YOKOHAMA MATHEMATICAL JOURNAL (YOKOHAMA)
1963/11 a 1977/25

ZEITSCHRIFT FÜR WAJSCHHEINLICHKEITSTHEORIE VERW. GEBIETE (BERLIN)
1962/1 a 1978/45
(FALTAM: vols. 34 a 36)

ZENTRALBLAT FÜR MATHEMATIK UND IHRE GRENZGEBIETE (BERLIN)
1967/135 a 1978/375
(FALTAM: vols. 172 a 178 e 190)

BIBLIOTECA DO CENTRO DE ESTUDOS DE ECONOMIA AGRARIA

BULLETIN (MENSUEL) D'ECONOMIE ET STATISTIQUES AGRICOLES (ROMA)
1956/5 a 1971/20

ECONOMETRICA (NEW HAVEN)
1944/12 a 1978/46

JOURNAL OF MATHEMATICAL ECONOMICS
1974/1 a 1978/5

REVIEW OF ECONOMICS AND STATISTICS (CAMBRIDGE)
1972/54 a 1978/60

ANÚNCIOS

CENTRO DE DOCUMENTAÇÃO CIENTÍFICA E TÉCNICA (CDCT) DO INSTITUTO NACIONAL DE INVESTIGAÇÃO CIENTÍFICA

O CDCT põe à disposição dos seus utilizadores os seguintes serviços modernos de informação e de documentação:

ALERTA PERMANENTE (SDI)

Basta definir o seu perfil científico ou técnico e receberá periodicamente listas de referências que o computador selecionará para si.

PESQUISA RETROSPECTIVA "ONLINE"

Em diálogo teleprocessado com as bases de dados localizadas na Europa ou nos Estados Unidos, um computador escolhe, para si, as referências bibliográficas que lhe interessam.

CATALOGAÇÃO

O serviço de catálogo pode localizar publicações periódicas existentes em Portugal, sobre qualquer assunto.

REPROGRAFIA

O CDCT pode obter para si cópias de documentos localizados quer no País quer no Estrangeiro.

ORIENTAÇÃO PARA AS FONTES DE INFORMAÇÃO

Recorrendo ao CDCT poderá saber onde deverá dirigir-se, no País ou no Estrangeiro, para obter a informação que deseja no caso de não a poder obter directamente.

CONTACTE O CDCT:

Av. Prof. Gama Pinto, 2 — 1600 LISBOA
Telefones: 77 28 86 - 73 13 00 - 73 13 50

BRASIL

Recebemos do Prof. Maurício de Pinho Gama, da Universidade de Brasília, Departamento de Estatística, a informação, com pedido de divulgação, da existência de vagas para pesquisadores/professores de Estatística nas Universidades do Brasil. Para mais detalhes contactar este Boletim.

7

Investigação (em) Estatística

CMA - Universidade Nova de Lisboa

Carlos Agra Coelho, Manuel L. Esquível, João Tiago Mexia,
cmac@fct.unl.pt, mle@fct.unl.pt, jtm@fct.unl.pt

*Centro de Matemática e Aplicações
Universidade Nova de Lisboa*

Estatística no CMA

1. Introdução

A fase inicial, aliás prolongada da história do CMA, pode ser considerada de maturação. Com efeito o CMA tem como instituição de acolhimento o Departamento de Matemática (DM) da FCT/UNL. Este departamento é jovem e apresenta uma grande preponderância do número de assistentes no seu corpo docente. Assim, nesta primeira fase, a ênfase, na área de Estatística, foi a preparação dos futuros membros doutorados do CMA.

Podemos considerar que a segunda fase se inicia em 2005. Nesta fase a ênfase foi posta na investigação aproximando-se o CMA do funcionamento de outros centros de investigação sediados em instituições, em que pondera o número de doutorados.

Assim, vamos centrar-nos nesta segunda fase. Presentemente o CMA tem três linhas de trabalho:

- Inferência Estatística;
- Teoria das Distribuições;
- Matemáticas Financeiras e Actuarias.

relacionadas com a Estatística.

No que se segue vamos considerar separadamente cada uma destas linhas concluindo com a apreciação global.

2. Inferência Estatística

Presentemente esta linha conta com os seguintes membros doutorados:

Nome	Ano de Doutoramento
João Tiago Mexia	1988
Sandra Nunes	2006
Luís Ramos	2007
João Lita da Silve	2007
Miguel Fonseca	2007
Ricardo Covas	2008
Vera de Jesus	2008

Os membros não doutorados são:

Elsa Moreira Gracinda Guerreiro Inês Sequeira Francisco Carvalho Miguel de Carvalho Paulo Canas Rodrigues
--

Vamos passar em revista os principais aspectos desenvolvidos pela equipa. Assim, temos:

2.1. Orientações de Doutoramento

Desde o início de 2005 foram aprovadas as seguintes dissertações de doutoramento orientadas ou co-orientadas por elementos da linha:

- Iola Pinto – "Joint Regression Analysis and Breeding Programs", 2005, UNL (orientador - J.T. Mexia)
- Ayana Furtado – "Modelação da estrutura dinâmica de povoamentos de *eucalyptus globulus* em primeira rotação", 2006, UNL (co-orientador - J.T. Mexia)
- Sandra Nunes – "Incidências – Modelo Logit e Medidas Aproximadas de Impactos Ambientais", 2006, UNL (orientador - J.T. Mexia)
- Dário Ferreira – "Variáveis pivot indutoras e componentes de variância em modelos normais ortogonais", 2006, UBI (orientador – J.T. Mexia)
- Sandra Ferreira – "Inferência para modelos ortogonais com segregação", 2006, UBI (orientador – J.T. Mexia)
- João Lita da Silva – "Consistência Forte de Estimadores: O Estimador dos Mínimos Quadrados", 2007, UNL (orientador – J.T. Mexia)
- Luís Ramos – "Quase-normalidade e inferência para séries de estudos emparelhados", 2007, UNL (orientador – J.T. Mexia)
- Vera M. Jesus – "Jordan algebras and crossing of factorial and fractional replicates", 2008, UNL (orientador – J.T. Mexia)
- Miguel Fonseca – "Inference and hypothesis testing in mixed linear models", 2007, UNL (co-orientador – J.T. Mexia)
- Ricardo Covas – "Orthogonal mixed models and commutative Jordan algebras", 2007, UNL (co-orientador – J.T. Mexia)
- Amílcar Oliveira – "Estabilidade em análise conjunta de regressões e condução dinâmica de planos de melhoramento", 2008, Univ. Aberta (orientador – J.T. Mexia)

Os elementos não doutorados encontram-se a preparar as respectivas dissertações:

Doutorando	Orientador
Elsa Moreira	João Tiago Mexia
Gracinda Guerreiro	João Tiago Mexia
Inês Sequeira	João Tiago Mexia
Francisco Carvalho	Manuela Oliveira (U. Évora)
Paulo Canas Rodrigues	Stanislaw Mejza (Poznan University of Life Sciences, Poland)

2.2. Conferências Internacionais

O responsável da linha – Doutor João Tiago Mexia, participou na organização das seguintes conferências:

- **SCRA 2006 – Statistics, Combinatorics and Related Areas**, Portugal Co-Chair da Comissão Organizadora Local Internacional
- **SCRA 2007 – Forum for Interdisciplinary Mathematics on Interdisciplinary Mathematical & Statistical Techniques**, China Membro da Comissão Científica
- **MAT-TRIAD 2007**, Polónia Membro da Comissão Científica
- **IMST 2007 – International Conference on Interdisciplinary Mathematical & Statistical Techniques**, China Membro da Comissão Científica Internacional

- **IWMS 2008 – International Workshop on Matrices and Statistics**, Portugal Chair da Comissão Organizadora

2.3. *Projectos I&D*

O CMA coordena um projecto sobre modelos lineares não ortogonais.

Em Portugal o projecto tem equipas a trabalhar no CMA e nos centros de investigação de Matemática da

- Universidade de Évora;
- Universidade da Beira Interior.

Na Polónia existe uma equipa integrada no projecto, no Instituto de Matemática da Universidade de Zielona Góra.

Além disso, membros da linha estão integrados em projectos sediados na Universidade Técnica de Lisboa (ISA):

- “Gestão de Secas”: Doutor João Tiago Mexia e Doutora Elsa Moreira;
- “Decision support tools for integrating fire and forest management planning”: Prof. Doutor João Tiago Mexia;

2.4. *Produção Científica*

Desde o início de 2005 foram publicados os seguintes trabalhos em revistas internacionais:

2008: 8 artigos (Zentralblat: 1; MathSciNet: 1; ISI: 3):

- Moreira, E.E.; Coelho, C.A.; Paulo, A.A.; Pereira, L.S.; Mexia, J.T. (2008). SPI-based drought category prediction using loglinear models. *Journal of Hydrology*. 354:116-130.
- Pereira, D. and Mexia, J.T. (2008). Selection proposal of cultivars of spring barley in the years from 2001 to 2004, using Joint Regression Analysis. *Plant Breeding*, 127(5): 452-458.
- Bras, A; Cotrim, CZ; Vasconcelos, I, et al. (2008). Asynchronous DNA replication detected by fluorescence in situ hybridisation as a possible indicator of genetic damage in human lymphocytes. *Oncology Reports*. 19(2):369-375.
- Ferreira, S.S., Ferreira, D. and Mexia, J.T. (2008). Double tier cross nesting design models. *Journal of Interdisciplinary Mathematics*. 11(2):275-289.
- Areia, A., Oliveira, M.M. and Mexia, J.T. (2008). Models for series of studies based on geometrical representation. *Statistical Methodology*. doi:10.1016/j.stamet.2007.09.001.
- Sequeira, I., Mexia, J.T. and Nunes, S. (2008). Double minimization for logit models with an additive two factors structure. *Biometrical Letters*, 45(1): 69-80.
- Oliveira, A., Oliveira, T. and Mexia, J.T. (2008). Analysis of residuals and adjustment in JRA. *Biometrical Letters*, 45(1): 45-54.
- Pereira, D. and Mexia, J.T. (2008). Application of Selective F tests in Joint Regression Analysis. *Journal of Statistical Theory and Practice*. 1(3-4):71-81.

2007: 16 artigos (Zentralblat: 2; MathSciNet: 9; ISI: 3):

- Fonseca, M.; Mathew, T.; Mexia, J.T.; Zmyslony, R. (2007). Tolerance intervals in the two-way nested model with mixed or random effects. *Statistics*, 41:289-300.
- Oliveira, M.M.; Mexia, J.T. (2007). ANOVA-like analysis of matched series of studies with a common structure. *Journal of Statistical Planning and Inference*, 137:1862-1870.
- Oliveira, M.M.; Mexia, J.T. (2007). Modelling series of studies with a common structure. *Computational Statistics and Data Analysis*, 51(12):5876-5885.
- Coelho, C.A.; Mexia, J.T. (2007). On the distribution of the product and ratio of independent generalized gamma ratio random variables. *Sankhya*, 65:221-255
- Jesus, V.; Rodrigues, P.C.; Mexia, J.T. (2007). Inference for random effects in prime basis factorials using commutative Jordan algebras. *Discussiones Mathematicae – Probability and Statistics*, 27:15-25

- Fernandes, C.; Ramos, P.; Ferreira, S.S.; Mexia, J.T. (2007). Variance components estimation in generalized orthogonal models. *Discussiones Mathematicae – Probability and Statistics*, 27:95-118
- Jesus, V.; Ferreira, S.S.; Mexia, J.T. (2008). Joint estimation for normal orthogonal mixed models. *Discussiones Mathematicae – Probability and Statistics*, 27:5-13
- Mexia, J.T.; Lita, J.L. (2008). Sufficient conditions for the strong consistency of least squares estimators with α -stable errors. *Discussiones Mathematicae – Probability and Statistics*, 27:27-45
- Ferreira, S.S.; Ferreira, D.; Mexia, J.T. (2007). Double tier cross-nesting design models. *Journal of Interdisciplinary Mathematics*, 11:278-285
- Fonseca, M.; Mexia, J.T.; Zmyslony, R. (2007). Jordan algebras, generating pivot variables and orthogonal normal models. *Journal of Interdisciplinary Mathematics*, 1:305-326
- Ferreira, S.S.; Ferreira, D.; Mexia, J.T. (2007). Cross additivity in balanced cross nesting models. *Journal of Statistical Theory and Practice*, 1:377-392
- Pereira, S.P., Oliveira, T.A. and Mexia J.T. (2007). Multiple Regression Models for Lactation Curves. *Biometrical Letters*. 44(1):71-81.
- Lima, A.T.; Fonseca, M.; Ribeiro, A.B.; Mexia, J.T.; Varela-Castro, A.; Ottosen, L. (2008). Modelling the effect of different conditions on electroremediation of fly ash from a municipal solid waste incinerator. *Biometrical Letters*, 44(1):51-70.
- Mexia, J.T. and Oliveira, M.M. (2007). Non-observable regressors. *Biometrical Letters*, 44(2):97-103
- Pereira, D.G., Mexia, J.T. and Rodrigues, P.C. (2007). Robustness of Joint Regression Analysis. *Biometrical Letters*, 44(2):105-128.
- Rodrigues, P.C. , Mexia, J.T. and Pereira, D.G. (2007). Analysis of Portuguese Life Tables. *Biometrical Letters*, 44(2):129-141.

2006: 9 artigos (Zentralblat: 5; MathSciNet: 6; ISI: 2):

- Fonseca, M., Mexia, J.T. and Zmyslony, R. (2006). Binary operations on Jordan algebras and orthogonal normal models. *Linear Algebra and its Applications*. 417: 75-86.
- Moreira, EE; Paulo, AA; Pereira, LS, et al. (2006). Analysis of SPI drought class transitions using loglinear models. *Journal of Hidrology*. 331(1-2):349-359.
- Nunes, C., Mexia, J.T. Non-central generalized F distributions. *Discussiones Mathematicae – Probability and Statistics*. 26(1):47-61.
- Mexia, J.T., Silva, J.L. (2006). Least squares estimator consistency: a geometric approach. *Discussiones Mathematicae – Probability and Statistics*. 26(1):19-45.
- Ferreira, S.S., Ferreira, D. and Mexia, J.T. (2006). Cross additivity - an application. *Discussiones Mathematicae – Probability and Statistics*. 26(2): 207-219.
- Nunes, C., Pinto, I. and Mexia, J.T. (2006). F and selective F tests with balanced cross-nesting and associated models. *Discussiones Mathematicae – Probability and Statistics*. 26(2):193-205.
- Rodrigues, P.C. and Mexia, J.T. (2006). ANOVA using commutative Jordan algebras, an application. *Discussiones Mathematicae – Probability and Statistics*. 26(2):179-191.
- Fernandes, C., Ramos, P. and Mexia, J.T. (2006). Evenness conditions for four-factor cross-nested models. *Biometrical letters*. 43(2):109-136.
- Sequeira, I. and Mexia, J.T. (2006). Permanent disability and social security. *Biometrical letters*. 43(2):1-8.

2005: 6 artigos (Zentralblat: 1; MathSciNet: 1; ISI: 1):

- Moreira, E.E., Ribeiro, A.B., Mateus, E.P. et al. (2005). Regressional modeling of electrolytic removal of Cu, Cr and As from CCA treated timber waste: application to sawdust. *Wood Science and Technology*. 39(4):291-309.
- Ferreira, D., Mexia, J.T., Ferreira, S.S. (2005). Stable hypothesis for mixed models with balanced cross-nesting. *Discussiones Mathematicae – Probability and Statistics*. 25(2):241-249.

- Fernandes, C. Ramos, P. and Mexia, J.T. (2005). Optimization of nested step designs. *Biometrical Letters*. 42(2):143-150.
- Ferreira, D. and Mexia, J.T. (2005). Confidence intervals for the variance components in a grapevine experiment. *Biometrical Letters*. 42(2):113-132.
- Fonseca, M., Jesus, V. and Mexia, J.T. (2005). Variance components estimation in nested designs. *Biometrical Letters*. 42(1):57-66.
- Moreira, E.E., Mexia, J.T., Ribeiro, A. Et al. (2005). Regression modeling of electrolytic removal of Cu, Cr and As from CCA treated timber waste: Application to wood chips. *Biometrical Letters*. 42(1):11-24.

2.5. *Apreciação sintética*

Pode concluir-se do que acima foi exposto que nesta fase, digamos de maturidade, a linha continua a desenvolver um trabalho continuado de preparação dos seus elementos.

O maior número de doutorados permitiu intensificar a produção científica. Nesta, temos a considerar uma vertente situada na área da Inferência Estatística Linear e que integra o projecto coordenado pelo CMA.

A par desta vertente teórica temos a das aplicações. Nesta vertente enquadra-se a participação nos dois projectos sediados na Universidade Técnica de Lisboa.

3. *Teoria das Distribuições*

Esta linha de investigação do CMA (Centro de Matemática e Aplicações da Universidade Nova de Lisboa) conta essencialmente com três áreas de investigação que são:

- Distribuições quase-exactas: desenvolvimento de distribuições quase-exactas para várias estatísticas, nomeadamente estatística de razão de verosimilhanças utilizadas em Estatística Multivariada;
- Distribuições exactas: obtenção de expressões manejáveis, com o potencial de uma utilização prática, para as distribuições exactas de várias estatísticas cujas expressões para a distribuição exacta não são conhecidas ou embora já conhecidas possuam expressões demasiado complicadas para que possam ter uma aplicação prática mesmo actualmente recorrendo a softwares potentes;
- Distribuições de extremos: investigação incidindo essencialmente sobre a estimação de parâmetros para distribuições de eventos raros.
- Modelos Lineares Generalizados: desenvolvimento de modelos lineares generalizados, nomeadamente modelos log-lineares, com vista à sua implementação em aplicações (principalmente estudos em hidrologia e linguística)

A equipa de investigação desta linha de investigação (subprojecto) é formada actualmente por 5 investigadores doutorados, 4 dos quais são ainda jovens investigadores, recém doutorados:

- Carlos Agra Coelho – doutorado em Bio-Estatística pela Universidade de Michigan em Maio de 1992 (com equivalência a Doutoramento em Estatística e Probabilidades atribuída pela Universidade de Lisboa em Outubro de 1992)
- Rui Pesado Alberto – doutorado em Matemática/Estatística pelo Instituto Superior de Agronomia da Universidade Técnica de Lisboa, em Maio de 2006
- Frederico Caeiro – doutorado em Estatística e Investigação Operacional pela Universidade de Lisboa em 2006
- Luís Miguel Grilo - doutorado em Matemática/Estatística pelo Instituto Superior de Agronomia da Universidade Técnica de Lisboa, em Julho de 2007
- Filipe José Marques – doutorado em Matemática, especialidade de Estatística pela Universidade Nova de Lisboa em Novembro de 2007.

Breve resumo das actividades desenvolvidas em cada uma das áreas de investigação na linha de Teoria das Distribuições

Distribuições quase-exactas

Esta é a área principal de investigação do responsável pela linha, tendo sido iniciada no fim dos anos 90, embora os primeiros artigos surjam somente no início dos anos 2000. Mais recentemente têm sido desenvolvidas distribuições quase-exactas para várias estatísticas de razão de verosimilhanças, nomeadamente de estatísticas de razão de verosimilhanças utilizadas em análise multivariada, como a estatística Lambda de Wilks, a estatística do teste de igualdade de várias matrizes de variância-covariância e a estatística do teste de esfericidade, utilizando vários tipos de decomposições das funções características do logaritmo destas estatísticas.

As distribuições quase-exactas assim obtidas apresentam não só um excelente comportamento assintótico em termos da dimensão das amostras, mas também em termos do número de variáveis envolvidas e, quando é o caso, também do número de blocos de variáveis e de amostras consideradas.

Ainda mais recentemente, foi também dado início ao desenvolvimento de distribuições quase-exactas para estatísticas utilizadas em testes de estruturas complexas para matrizes de variância-covariância, como as correspondentes aos testes de esfericidade matricial para uma ou várias populações, esfericidade por blocos para uma ou várias populações e esfericidade múltipla para várias populações. As distribuições quase-exactas para as estatísticas destes testes serão obtidas a partir da decomposição das respectivas hipóteses nulas em sequências de hipóteses elementares encaixadas, condicionalmente independentes, permitindo assim a construção de funções características quase-exactas a partir da decomposição das funções características das estatísticas correspondentes às hipóteses elementares.

A comparação dos resultados obtidos com aproximações assintóticas e de ponto-de-sela mostram não só a extrema proximidade das distribuições quase-exactas obtidas às respectivas distribuições exactas como também nos dão excelentes expectativas para trabalhos futuros.

A grande manejabilidade das distribuições obtidas permite também um fácil cálculo de valores-de-p e quantis quase-exactos, os quais, dada a sua qualidade, podem ser utilizados em substituição dos exactos.

Fazendo a ponte para a área das distribuições exactas, apresentada a seguir, e também para a área das distribuições de extremos, encontram-se em curso vários trabalhos onde se estende a várias outras estatísticas, como por exemplo a estatísticas F generalizadas, à variância generalizada e estatísticas cuja distribuição é a de um produto de variáveis aleatórias Gama independentes, ou da soma de variáveis aleatórias Gumbel generalizadas independentes, o desenvolvimento de distribuições quase-exactas, uma vez que as distribuições exactas destas estatísticas não são manejáveis e as usuais aproximações assintóticas nem sempre apresentam a qualidade desejada. Alguns destes trabalhos estão a ser desenvolvidos em colaboração com outros membros de outras linhas do CMA, como o Professor João Tiago Mexia, Coordenador Científico do CMA e responsável pela linha de investigação (sub-projecto) em Inferência Estatística e com o Professor Barry C. Arnold do Departamento de Estatística da Universidade da Califórnia em Riverside.

Distribuições exactas

Nesta área de investigação têm sido desenvolvidos estudos sobre a obtenção de formas e expressões manejáveis para as distribuições exactas de várias estatísticas como sejam as estatísticas F generalizadas, a variância generalizada e outras estatísticas cuja distribuição é a de um produto de variáveis aleatórias Gama independentes. Pretende-se também que tais formas das distribuições exactas possam ser utilizadas no desenvolvimento de distribuições quase-exactas para as mesmas estatísticas.

Nesta área de investigação também se tem ainda desenvolvido investigação conducente ao desenvolvimento de distribuições multivariadas com distribuições marginais Gama Inteira Generalizadas e de distribuições Gama, Gama Inteira Generalizada, Laplace, e outras relacionadas com a distribuição Gama, sobre a circunferência (as chamadas distribuições circulares ou 'enroladas').

Distribuições de Extremos

O trabalho de investigação desenvolvido nesta área desenvolve-se sobretudo em torno da estimação não-paramétrica de parâmetros em modelos para eventos raros, em colaboração com a Professora Ivette Gomes do Departamento de Estatística e Investigação Operacional da Faculdade de Ciências da Universidade de Lisboa.

Modelos Lineares Generalizados

Esta é de facto uma área menor de investigação dentro da linha (sub-projecto) de Teoria das Distribuições, cujas actividades se relacionam sobretudo com o desenvolvimento de Modelos Lineares Generalizados, nomeadamente Modelos Log-lineares, com vista à sua aplicação em estudos de Linguística e Hidrologia. Especial atenção é dedicada a estudos sobre as distribuições assintóticas dos parâmetros e funções de parâmetros nestes modelos, como sejam os estimadores de razões de *odds* nos modelos Log-lineares utilizados.

3.1. Orientações de Doutoramento

- Alberto, R. P. (2006). *Desenvolvimento de distribuições quase-exactas para várias estatísticas de razão de verosimilhanças utilizadas em Estatística Multivariada, nomeadamente a estatística Lambda de Wilks*. Tese de Doutoramento em Matemática/Estatística, Instituto Superior de Agronomia, Universidade Técnica de Lisboa. (orientador: Carlos A. Coelho)
- Caeiro, F. (2006). *Estimação de parâmetros em eventos raros*. Tese de Doutoramento em Estatística e Investigação Operacional, Faculdade de Ciências, Universidade de Lisboa. (orientador: M. Ivette Gomes)
- Grilo, L. M. (2007). *Desenvolvimento de distribuições quase-exactas para vários cenários de utilização da estatística Lambda de Wilks*. Tese de Doutoramento em Matemática/Estatística, Instituto Superior de Agronomia, Universidade Técnica de Lisboa. (orientador: Carlos A. Coelho)
- Marques, F. J. (2007). *Distribuições quase-exactas para estatísticas de razão de verosimilhanças*. Tese de Doutoramento em Matemática, especialidade de Estatística, Universidade Nova de Lisboa. (orientador: Carlos A. Coelho; co-orientador: João T. Mexia)

3.2. Conferências Internacionais

Dois dos membros desta linha de investigação fizeram parte da Comissão Organizadora da Conferência SCRA2006/FIM XIII – International Conference on Interdisciplinary Mathematical & Statistical Techniques and XIIth Conference of the Fórum for Interdisciplinary Mathematics que decorreu em Tomar de 1 a 4 de Setembro de 2006, Carlos A. Coelho como Presidente da Comissão Organizadora Local e Luís M. Grilo como Coordenador Local, tendo ainda também o Professor João Tiago Mexia integrado a Comissão Organizadora Local. No âmbito desta Conferência foi atribuído o grau de Doutor Honoris Causa ao Professor C. R. Rao. A Conferência contou com a apresentação de mais de 270 comunicações e com a participação de mais de 300 Conferencistas vindos de vários países.

Carlos A. Coelho tem também integrado as Comissões Organizadoras e Científicas de várias Conferências nacionais e internacionais, como por exemplo:

- Membro da Comissão Organizadora Local do *IWMS 08 – 17th International Workshop on Matrices and Statistics*, que decorreu em Tomar, de 23 a 26 de Julho de 2008, cujo Presidente da Comissão Organizadora foi o Professor João Tiago Mexia
- Membro do Scientific Committee do *3rd Workshop on Statistics, Mathematics and Computation, 1st Portuguese-Polish Workshop on Biometry* da Universidade Aberta que decorreu de 21 a 22 de Julho de 2008
- Membro do FIM's (Forum for Interdisciplinary Mathematics) Conference Organizing Committee da *International Conference on Interdisciplinary Mathematical & Statistical Techniques, XVI-th Conference of the Forum for Interdisciplinary Mathematics (IMST'2008/FIM XVI)* que decorreu de 16 a 18 de Maio de 2008 na Universidade de Memphis, Tennessee, EUA.
- Membro da Comissão Científica (International Advisory Committee) da *International Conference on Advances in Interdisciplinary Statistics and Combinatorics*, realizada de 12 a 14 de Outubro de 2007 na Universidade da Carolina do Norte, E.U.A.
- Membro da Comissão Científica do *II Workshop em Estatística, Matemática e Computação da Universidade Aberta*, 8-9 de Novembro de 2006.
- Membro da Comissão Científica (International Advisory Committee) do *SCRA 2004/FIM XI (Eleventh International Conference on Interdisciplinary Mathematical and Statistical Techniques)* realizado de 27 a 29 de Dezembro de 2004 no Institute of Engineering and Technology de Lucknow, Índia.

- Membro da Comissão Científica (International Advisory Committee) do *SCRA 2003/FIM X (International Conference on Statistics, Combinatorics and Related Areas, 10th Conference of the Forum for Interdisciplinary Mathematics)* realizado de 3 a 5 de Outubro de 2003 na University of Southern Maine, Portland, Maine, USA.

3.3. *Projectos I&D*

Carlos Coelho participa ou participou em:

- PTDC/AGR-AAM/71649/2006 -- Gestão do risco em secas: identificação, monitorização, caracterização, predição e mitigação' (2007-2009), da responsabilidade do Prof. Luís Santos Pereira do Instituto Superior de Agronomia/UTL.
- Projecto SEDEMED II - "Secheresse et Désertification dans le bassin Méditerranéen II" que se enquadra no âmbito do programa comunitário Interreg III B Medocc (2003-2004), cujo responsável em Portugal foi o Prof. Luís Santos Pereira do Instituto Superior de Agronomia/UTL.
- AGRO~508 - 'Compostagem de Resíduos Agro-Industriais e sua valorização na fertilização de culturas hortícolas e olival em Agricultura Biológica' (2004-2006), da responsabilidade da Prof.^a Ana Cristina Queda do ISA/UTL;
- AGRO~287 - 'Avaliação do efeito da isenção de viroses em pereiras e macieiras portuguesas e fornecimento de material vegetativo com garantia sanitária' (2001-2004), da responsabilidade da Prof.^a Maria Teresa Barros do ISA/UTL;
- AGRO~060 -- 'Práticas de controlo da dormência para a regularização das colheitas em pereira, macieira e ameixiera' (2001-2004), da responsabilidade do Prof. Nuno Barba da Escola Sup. Agrária de Santarém.

Frederico Caeiro participa ou participou em:

- Computational Statistical and Applications, Projecto no1141, do Centro de Matemática e Aplicações (2003-2006).
- ERAS - Extremes, Risk, Safety and the Environment (ERSE), POCI/MAT/58876/2004 (2005-2008).

3.4. *Produção Científica (Zentralblatt Math e/ou MathSciNet) (2005-2008:17)*

- Caeiro, F., Gomes, M. I. and Pestana, D. (2005). Direct Reduction of Bias of the Classical Hill Estimator. *Revstat*, **3**, 113-136.
- Caeiro, F., Gomes, M.I. (2006). A new class of estimators of a "scale" second order parameter". *Extremes*, **9**, 193-211.
- Coelho, C. A. (2006). The exact and near-exact distributions of the product of independent Beta random variables whose second parameter is rational. *Journal of Combinatorics, Information & System Sciences*, **31**, 11-34. [MR2351709]
- Coelho, C. A., Alberto, R. P., Grilo, L. M. (2006). A mixture of Generalized Integer Gamma distributions as the exact distribution of the product of an odd number of independent Beta random variables. *Journal of Interdisciplinary Mathematics*, **9**, 229-248. [MR2245158] [Zbl 1117.62017]
- Coelho, C. A., Alberto, R. P., Grilo, L. M. (2006). A mixture of Generalized Integer Gamma distributions as the exact distribution of the product of an odd number of independent Beta random variables, in *Advances in Interdisciplinary Mathematics*, S. Gupta & B. K. Dass (eds.), 1-20 (reimpressão de *J. of Interdisciplinary Mathematics*)
- Coelho, C. A. (2006). The joint characterization of discrete and continuous 'waiting times' through their reciprocal relationship. *Journal of Interdisciplinary Mathematics*, **9**, 297-318. [Zbl pre05187067]
- Coelho, C. A. (2006). The joint characterization of discrete and continuous 'waiting times' through their reciprocal relationship, in *Advances in Interdisciplinary Mathematics*, S. Gupta & B. K. Dass (eds.), 69-90 (reimpressão de *J. Interdisciplinary Mathematics*).
- Alberto, R. P., Coelho, C. A. (2007). Study of the quality of several asymptotic and near-exact approximations based on moments for the distribution of the Wilks Lambda statistic. *Journal of Statistical Planning and Inference*, **137**, 1612-1626. [MR2303779] [Zbl 1110.62021]

- Grilo, L. M., Coelho, C. A. (2007). Development and study of two near-exact approximations to the distribution of the product of an odd number of independent Beta random variables. *Journal of Statistical Planning and Inference*, **137**, 1560-1575. [MR2303776] [Zbl 1110.62020]
- Coelho, C. A., Mexia, J. T. (2007). On the distribution of the product and ratio of independent generalized gamma-ratio random variables. *Sankhya*, **69**, 2, 221-255. [MR2428869]
- Coelho, C. A. (2007). The wrapped Gamma distribution and wrapped sums and linear combinations of independent Gamma and Laplace distributions. *Journal of Statistical Theory and Practice*, **1**, 1-29. [MR2354613] [Zbl pre05235904]
- Marques, F. J., Coelho, C. A. (2008). Near-exact distributions for the sphericity likelihood ratio test statistic. *Journal of Statistical Planning and Inference*, **138**, 726-741. [MR2382885 (2008k:62115)]
- Caeiro, F. and Gomes, M.I. (2008). Minimum-Variance Reduced-Bias Tail Index and High Quantile Estimation. *Revstat*, **6**, 1-20.
- Coelho, C. A., Marques, F. J. (2008). The advantage of decomposing elaborate hypotheses on covariance matrices into conditionally independent hypotheses in building near-exact distributions for the test statistics. *Linear Algebra and Its Applications* (a aguardar publicação).
- Caeiro, F. and Gomes, M.I. (2008). Semi-parametric second order reduced-bias high quantile estimation. *Test* (a aguardar publicação).
- Gomes, M.I., Pestana, D. and Caeiro F. (2008). A note on the asymptotic variance at optimal levels of a bias-corrected Hill estimator. *Statistics and Probability Letters* (a aguardar publicação).
- Caeiro, F., Gomes, M.I. and Henriques Rodrigues, L. (2008). Reduced-bias tail index estimators under a third order framework. *Communications in Statistics - Theory and Methods*. (a aguardar publicação).

4. Matemáticas Actuariais e Financeiras

A linha de Matemáticas Actuariais e Financeiras cobre um vasto leque de assuntos que passamos a apresentar brevemente associando-os aos investigadores que os desenvolvem.

Manuel Leote Tavares Inglês Esquível, doutorado em 1997, em Processos Estocásticos, na FCT/UNL. Desenvolvem-se métodos para avaliação do risco de investimentos baseados em dados de preço/liquidez. Estudou-se um modelo para a evolução acoplada dos preços Spot/Futuros de matérias primas. Estudaram-se algoritmos aleatórios para optimização estocástica global. Estudaram-se as funções geradoras de probabilidade para variáveis aleatórias discretas tomando valores reais.

Marta Cristina Vieira Faias Mateus, doutorada em 2000, em Economia Matemática, na FE/UNL. No contexto da teoria geral do equilíbrio providencia-se um modelo de mercados incompletos em que restrições privadas no consumo são interpretadas com falta de informação. Prova-se a existência de equilíbrio quando os agentes são incapazes de inferir qualquer informação adicional a partir dos preços. Seguidamente, estuda-se o grau de indeterminação para este conjunto de equilíbrio. No contexto de formação de clubes explora-se a existência de equilíbrios Bayesianos para uma economia com informação incompleta sobre tipos inerentes.

Rui Manuel Rodrigues Cardoso. Em 2004, obteve o PhD in Actuarial Mathematics and Statistics, School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, Scotland, UK. Cálculo de aproximações e majorantes da probabilidade de ruína em horizonte finito e tempo contínuo, considerando diversas generalizações do modelo clássico de risco, e obtenção também de majorantes e aproximações do valor esperado descontado dos dividendos pagos aos accionistas considerando modificações do modelo clássico de risco, como por exemplo a introdução de uma barreira. A ferramenta usada para a obtenção destas quantidades são as cadeias de Markov, tendo por base os artigos Cardoso and Waters (2003) *Recursive calculation of finite time ruin probabilities under interest force* e Cardoso and Waters (2005) *Calculation of finite time ruin probabilities for some risk models*, ambos publicados na revista *Insurance: Mathematics and Economics*,

Maria de Lourdes Belchior Afonso, doutorada em 2008, no ISEG/UTL, em Matemática Aplicada à Economia e à Gestão. Obteve-se um método para o cálculo da probabilidade de ruína em tempo contínuo e horizonte finito para um processo de Poisson composto onde o prémio é constante ao longo de cada período de tempo (ano), mas depende da informação passada de indemnizações agregadas anuais. Em função disso, o prémio é ajustado anualmente, passando a ser variável de período para

período. A publicação *Calculating continuous time ruin probabilities for a large portfolio with varying premiums* submetida ao ASTIN Bulletin foi aceite em 23 Dezembro 2008

Pedro Palhinhas Mota, obteve em 2008 o doutoramento em Estatística na FCT/UNL orientado pelos Professores Doutores Mathieu Kessler, da Universidade de Cartagena, Espanha, e Manuel L. Esquível, FCT, Universidade Nova de Lisboa, na área da estimação de limiares em processos contínuos com regimes.

Miguel Brás de Carvalho, prepara-se para submeter a sua dissertação de doutoramento para provas públicas na área dos algoritmos estocásticos para provas públicas. Os trabalhos decorreram sob a orientação do Professores João Tiago Mexia e Manuel L. Esquível.

Carlos Manuel Antunes Veiga, encontra-se actualmente a frequentar o programa de PhD na Frankfurt School of Finance & Management, Centre for Practical Quantitative Finance, sob a direcção do Prof. Dr. Uwe Wystup. Trabalha no campo da Matemática Financeira, focado em temas relevantes para o apreçamento e cobertura de risco de contratos financeiros. Em particular, investigação no tema do apreçamento de opções exóticas sobre acções, e da cobertura de risco de opções em modelos que incluem dividendos discretos. Apresentação numa conferência internacional e apresentação, SCRA2006, e num colóquio dedicado à matemática financeira, Frankfurt MathFinance Colloquium.

Gonçalo dos Reis aguarda as provas públicas de doutoramento sob a orientação de P. Imkeller. A dissertação lida com métodos numéricos para equações diferenciais estocásticas retrógradas com termos de crescimento quadrático. A publicação *Pricing and hedging of derivatives based on non-tradable underlyings*, aparecerá na revista *Mathematica Finance*. Uma outra publicação *On Securitization, Market Completion and Equilibrium Risk Transfer*, está submetida para publicação. Encontra-se actualmente na École Normale Supérieure de Paris com uma bolsa trabalhando com o Professor Nizar Touzi.

João Beleza de Sousa, colaborador do CMA, está a iniciar os trabalhos de doutoramento na área dos processos Gaussianos na óptica do ‘machine learning’.

A Doutora Alexandra Dias, em 2005, continuou a sua estadia como Research-Fellow no RiskLab, ETH Zurich, sob orientação do Professor Doutor Paul Embrechts. O trabalho consistiu em aprofundar o estudo de detecção de pontos de mudança na estrutura de dependência de séries temporais, aplicado a taxas de câmbio (FX). Em 31 de Julho, a Doutora Alexandra Dias desvinculou-se do CMA.

De entre as publicações nacionais produzidas nestes trabalhos salientamos as seguintes.

Esquível, M. L.; Mexia, J. T.; Corte Real P.; Lita da Silva, J.; Convergência Forte de Estimadores, Erros com Decaimento Exponencial no Infinito e Colapso de Erros Radiaes Actas do XIII Congresso Anual da SPE 2006.

Esquível, M. L.; Mota, P. P. Sobre a Estimação de Limiares em Processos Contínuos com Regimes, Actas do XIII Congresso Anual da SPE 2006.

Afonso, M. L.; Mexia, J. T.; Estratégias Óptimas de Resseguro- Excess of Loss, Boletim do Instituto Português de Actuários nº 41 2006.

Foram também produzidos os seguintes relatórios técnicos:

Faias, M; Information and financial Markets Pré-Publicações Nº 13/16 DMat-FCT-UNL

Faias, M; A Strategic Club Formation; Existence and Characterization of equilibrium Pré-Publicações Nº 14/16 DMat-FCT-UNL

4.1. *Orientações de Doutoramento*

No período 2006-2008 realizaram-se as orientações de doutoramento de Maria de Lourdes Afonso, Pedro Palhinhas Mota e Miguel Brás de Carvalho e Gonçalo dos Reis. Estão em curso as orientações de Carlos Antunes Veiga e Miguel Beleza.

4.2. *Conferências Internacionais*

Marta Faias, Consumption Constraints and Informational Selections of Financial Equilibria, XVI European Workshop on General Equilibrium Theory Warwick University (UK). June, 2007.

Marta Faias, Consumption Constraints and Informational Selections of Financial Equilibria, 8th International Conference on Current Trends in Economics Society for the Advancement of Economic Theory (Kos (Grécia)).

Marta Faias, RGEA Workshop on Microeconomics and Mathematical Economics, Universidad de Vigo, Vigo, Espanha, Outubro, 2007.

Maria de Lourdes Afonso, IME 2007 – Julho de 2007 – Pireaus, Greece. Presentation: Continuous time ruin probabilities for a portfolio with credibility-adjusted premiums. Afonso, L.B.; Egídio dos Reis, A; Waters, H.

Rui Cardoso - "Dividends and ruin problems in finite time", Eleventh International Congress on Insurance: Mathematics and Economics, University of Pireus, Greece, July 11, 2007

Manuel L. Esquível, Workshop "Innovations in Mathematical Finance" at Loen (Norway) between June 25th and July 1st 2007, organized by the AMAMEF

Manuel L. Esquível, International Conference on Advances in Interdisciplinary Statistics and Combinatorics, October 12-14, 2007 – Greensborough, USA. Presentation: "On some stochastic algorithms for global optimization".

Carlos Veiga, International Conference on Advances in Interdisciplinary Statistics and Combinatorics, October 12-14, 2007 – Greensborough, USA. Presentation: "European Options with Discrete Dividends".

Miguel Carvalho, An Approach to Stochastic Optimization, Advances in Interdisciplinary Statistics and Combinatorics, University of North Carolina - US.

Miguel Carvalho, Forecasting the Unemployment Rate in the Euro Area, Poster Presentation, Euroconference Series in Quantitative Economics and Econometrics: (EC)2, Universidade do Algarve - Faculdade de Economia - Portugal.

Gonçalo dos Reis, UBC Probability Seminars at the University of British Columbia (UBC) in Vancouver (Canada), March 21, 2007,

Gonçalo dos Reis, Classical and Variational Differentiability of BSDEs with Quadratic Growth, International Research Trainings Group Seminar at WIAS in Berlin (Germany), May 16, 2007,

Gonçalo dos Reis, Classical and Variational Differentiability of BSDEs with Quadratic Growth, Workshop Innovations in Mathematical Finance organized by the Norway CMA held in Loen (Norway), June 26th to July 1st, 2007

Gonçalo dos Reis, Differentiability of quadratic BSDEs and applications to optimal hedging of derivatives on non-tradable underlyings, Conference 56th session of the ISI organized by the International Statistical Society, held in Lisbon (Portugal) from the 22nd to the 29th August, 2007

Gonçalo dos Reis, Classical and Variational Differentiability of BSDEs with Quadratic Growth Conference 8th German Open Conference on Probability and Statistics, held in Aachen (Germany) from the 4th to the 7th March, 2008

Gonçalo dos Reis, Pricing and hedging of derivatives based in non-tradable underlyings Conference Bachelier Finance Society: Fifth World Congress, held in London (England) from the 15th to the 19th July, 2008

Carlos Veiga, International Conference on Advances in Interdisciplinary Statistics and Combinatorics, October 12-14, 2007 – Greensborough, USA. European Options with Discrete Dividends

Carlos Veiga, SCRA2006, September 2006 – Lisbon European Options with Discrete Dividends - An Hedging Perspective.

Carlos Veiga, MathFinance Colloquium, June 2006 – Frankfurt am Main Exploring the Limits of Closed Pricing Formulas in the Black and Scholes Framework.

Carlos Veiga, Stochastic Finance, September 2004 – Lisbon Expanding further the Universe of Exotic Options Closed Pricing Formulas in the Black and Scholes Framework.

Carlos Veiga, Congress of the Bachelier Finance Society, July 2004 – Chicago Expanding the Universe of Exotic Options Closed Pricing Formulas in the Black and Scholes Framework.

4.3. *Projectos I&D*

Marta Faias - SA070A05: Junta de Castilla y Leon. "Informacion, Coaliciones y Manipulacion". 5.733 Euros

Marta Faias - PTDC/ECO/64968/2006: Fundação para a Ciência e Tecnologia, "Crenças, especulação e equilíbrio temporário, Beliefs, speculation and temporary equilibrium".

Manuel L. Esquível - projecto de investigação com o Banco de Investimento Global denominado "Pricing of Interest rate derivatives" com o financiamento de 9600 Euros.

4.4. *Produção Científica* (Zentralblatt Math e/ou MathSciNet e/ou ISI) (2005-2008)

Esquível, M. L.; Probability generating functions for discrete real-valued random variables. *Teor. Veroyatn. Primen.* 52 (2007), no. 1, 129–149; *translation in Theory Probab. Appl.* 52 (2008), no. 1, 40–57.

Esquível, M. L.; Grossinho, M. R.; Shiryaev, A.; Oliveira, P. E. (editores) ; *Stochastic Finance* Springer 2006, 364 páginas.

Esquível, M. L.; A Conditional Gaussian Martingale Algorithm for Global Optimization *Lecture Notes in Computer Science* 3982 841-851.

Ankirchner, Stefan; Imkeller, Peter; Dos Reis, Gonçalo; Classical and variational differentiability of BSDEs with quadratic growth. *Electron. J. Probab.* 12 (2007), no. 53, 1418–1453 (*electronic*).

Faias, Marta; Approximate equilibrium in pure strategies for a two-stage game of asset creation. *Decis. Econ. Finance* 31 (2008), no. 2, 117–136.

Cardoso, Rui M. R. ; Waters, Howard R.; Calculation of finite time ruin probabilities for some risk models. *Insurance Math. Econom.* 37 (2005), no. 2, 197–215.

5. *Observações finais*

Do atrás exposto conclui-se facilmente o grande desenvolvimento da Estatística e disciplinas afins no CMA.

Saliente-se, a par da investigação o esforço feito para apoiar alunos de doutoramento.



GPE - Universidade de Aveiro

Manuel Scotto, *mscotto@ua.pt*

*Grupo de Probabilidades e Estatística
Departamento de Matemática
Universidade de Aveiro*

O Grupo de Probabilidades e Estatística (GPE) faz parte da Unidade de Investigação ‘Matemática e Aplicações’ sediada no Departamento de Matemática da Universidade de Aveiro (UA). O objectivo fundamental do GPE é desenvolver investigação em Probabilidades e Estatística Matemática assim como contribuir tanto para a consolidação do trabalho e maturação científica de jovens investigadores, como para a formação de novos investigadores através da formação pós-graduada. Na actualidade o GPE é constituído por 8 membros efectivos e 6 colaboradores.

As principais linhas de investigação do GPE centram-se na teoria de valores extremos, na robustez, em séries temporais, em bioestatística e nas aplicações da Estatística em áreas como meio ambiente, engenharia e medicina. Na área de valores extremos tem sido dado especial ênfase ao estudo das características extremas associadas a (a) sucessões estacionárias sub-amostradas; (b) sucessões desagregadas; (c) sucessões de contagem; (d) equações estocásticas às diferenças; (e) distribuições max-semi-estáveis; e finalmente às sucessões de variáveis duplamente indexadas (esquema triangular). Outro tópico em destaque é o problema da escolha estatística dos domínios de atracção para valores extremos. Várias estatísticas de teste foram propostas tendo-se revelado muito convenientes na discriminação de distribuições de caudas pesadas e super pesadas. Na área da robustez estatística têm sido investigados procedimentos que permitam aplicar os princípios fundamentais da robustez a alguns modelos estatísticos com estruturas de dependência específicas, nomeadamente, a modelos de equações simultâneas e a modelos geoestatísticos. Os modelos de equações simultâneas são úteis, sobretudo, na descrição do relacionamento entre variáveis macro-económicas; apresentam uma caracterização funcional que sugere a utilização de técnicas da regressão robusta combinadas com métodos de mínimos quadrados por fases ou com o método de momentos generalizados. A modelação geoestatística, que foi inicialmente desenvolvida tendo em vista a descrição de fenómenos geológicos, é actualmente usada no estudo de características espaciais de natureza diversa, como por exemplo, em aplicações epidemiológicas ou ambientais. No que diz respeito a esses modelos, os tópicos analisados incidem na estimação robusta do variograma. A estrutura de dependência traduzida pela função de variograma exige uma combinação especialmente delicada entre princípios de robustez e de eficiência, pois requer que os métodos robustos apresentem bom desempenho com poucas observações; para além disso, envolve uma componente considerável de estatística computacional. Outro tópico em investigação sobre robustez estatística incide no desenvolvimento de técnicas de diagnóstico baseadas em estatísticas robustas, para aplicação em modelos de regressão ou em modelos com erros-nas-variáveis. Outra área com especial relevância no GPE são as séries temporais. Neste caso o destaque tem sido dados à caracterização e inferência (do ponto de vista clássico e Bayesiano) de modelos não lineares, nomeadamente do tipo bilinear, modelos autoregressivos e médias móveis para sucessões de contagem e modelos autoregressivos por limiares. Outra área de relevância dentro das séries temporais que tem vindo a ser analisada por membros do GPE, é a implementação de sistemas de alarme óptimos para processos de contagem. Na área da bioestatística o GPE colabora com investigadores do Departamento de Electrónica e Telecomunicações e do Departamento de Biologia (formando o grupo de Bioinformática) da UA. Tem com principal objectivo o desenvolvendo de metodologias estatística para análise de sequências genómicas (núcleótidos, codões e aminoácidos) e de dados de microarrays, que auxiliem o investigador na identificação, caso existam, de propriedades estatísticas relevantes nas sequências que regulam o código genético de qualquer espécie codificada e na identificação de padrões ou grupos de genes diferencialmente expressos em matrizes de dados de microarrays.

Finalmente, as aplicações da Estatística têm-se centrado principalmente no estudo de dados biomédicos, meteorológicos (índices de pluviosidade), oceanográficos (altura significativa de onda) e ambientais (níveis de ozono). Outro assunto que tem merecido especial destaque é o estudo da Esclerose Lateral Amiotrófica (ELA) em colaboração com o Hospital de Santa Maria, o Centro de Estudos Egas Moniz, o Centro de Medicina de Reabilitação da Região Centro (Tocha) e o Royal London Hospital and Queen Mary's School of Medicine. A ELA é uma doença neuromuscular do neurónio motor que leva a uma atrofia muscular progressiva com correspondente perda de força muscular. A ELA é uma doença de evolução muito rápida, sendo o tempo médio de vida 3 anos após o início da doença. Um dos problemas que se colocam no estudo da ELA é a necessidade de identificar parâmetros neurofisiológicos ou clínicos sensíveis para prever a progressão da doença. Um novo indicador para avaliar a progressão da ELA foi proposto tendo-se revelado um parâmetro útil para a previsão da evolução da doença.

Em síntese, os tópicos de investigação de maior relevância do grupo de Probabilidades e Estatística concentram-se na teoria de valores extremos, séries temporais não lineares, robustez e genética. De acordo com o relatório da FCT 2007, um dos pontos fortes deste grupo consiste na aplicação da estatística a problemas de modelação ambientais, biologia e genética, devendo-se realçar a existência de vários projectos interdisciplinares.



CIMA - Universidade de Évora

Carlos Braumann, *braumann@uevora.pt*

Russell Alpizar-Jara, *alpizar@uevora.pt*

*Centro de Investigação em Matemática e Aplicações
Universidade de Évora*

O Centro e os seus Objectivos

O Centro de Investigação em Matemática e Aplicações (CIMA-U.E.) foi criado em 1993, regendo-se actualmente pelo Regulamento constante da Ordem de Serviço 4/2005 da Universidade de Évora. É presentemente o único centro de investigação em Matemática na parte sul de Portugal Continental. O CIMA-U.E. tem como principais objectivos:

- a) Promover a investigação em Matemática, particularmente a suscitada pelas aplicações.
- b) Apoiar a realização de acções de formação de recursos humanos no seu domínio.
- c) Difundir o conhecimento científico na sua área, nomeadamente através da edição de publicações e da realização de encontros, colóquios e congressos.
- d) Promover o intercâmbio científico com instituições e investigadores da mesma área e áreas afins.
- e) Contribuir, no seu domínio, para o desenvolvimento e a modernização dos sectores produtivos, regionais e nacionais.

O CIMA-U.E. é uma unidade de I&D financiada pela Fundação para a Ciência e Tecnologia (FCT), através do Programa Financiamento Plurianual, que tem vindo a crescer e expandir as suas áreas de acção. Actualmente as actividades do CIMA-U.E. organizam-se por projectos de investigação e o Centro conta com 31 investigadores doutorados e 18 estudantes de doutoramento afectos essencialmente a 3 projectos de investigação:

1. Equações Diferenciais e Optimização (13 doutorados e 6 doutorandos). Alguns membros deste projecto também fazem investigação em Análise Funcional, Análise Não-suave, Cálculo das Variações, Análise Numérica.
2. Sistemas Dinâmicos e Estruturas Discretas (12 doutorados e 6 doutorandos). Análise Não Standard, Álgebra e Geometria estão dentre as áreas de interesse dos membros deste projecto.
3. Aplicações dos Processos Estocásticos e Estatística às Ciências da Vida (6 doutorados e 6 doutorandos).

Estes projectos resultaram duma recente reestruturação, seguindo as linhas orientadoras dos Painéis de Avaliação Internacional nomeados pela FCT, que reconheceram a importância do Centro focar, nas aplicações, a área da Biomatemática. De facto, os membros do último projecto, que desenvolvem investigação nas aplicações da Estatística e Processos Estocásticos às Ciências da Vida, têm também apostado nessa área de investigação. Tratando-se do Boletim da SPE, interessa certamente a investigação do Centro na área da Estatística, pelo que focaremos a seguir os trabalhos desenvolvidos no âmbito do terceiro projecto da lista.

O Projecto *Aplicações dos Processos Estocásticos e Estatística às Ciências da Vida* ou *Stochastic and Statistical Applications in Life Sciences* (SSALS)

Este projecto tem como principal objectivo o desenvolvimento e a aplicação de modelos estatísticos em Ciências da Vida. Interessam, por exemplo, as aplicações nas áreas da Biologia, Epidemiologia, Agricultura e Gestão dos Recursos Naturais, entre outras. São membros do projecto os doutorados Carlos Braumann (coordenador), Dulce Gamito, Dulce Gomes, Manuela Oliveira, Paulo Infante e Russell Alpizar-Jara, e os estudantes de doutoramento Anabela Afonso, Gonçalo Jacinto, Inês Sousa Dias, João Monteiro, Manuel Minhoto e Patrícia Filipe. Espera-se que estes doutorandos completem as suas teses brevemente.

Os desenvolvimentos teóricos relativos a este projecto são especialmente motivados pela necessidade de resolver problemas aplicados. O projecto consta de 5 tarefas com interacções entre algumas delas:

Tarefa 1: Crescimento de populações e pesca em ambiente aleatório

O objectivo com esta tarefa é o desenvolvimento, extensão e aplicação de novos modelos de crescimento de populações animais em ambientes determinístico e estocástico. Interessa a generalização de Equações Diferenciais Estocásticas (EDE) em modelos de crescimento populacional, tempos de extinção, existência de densidades estacionárias e a resolução da controvérsia dos cálculos Itô/Stratonovich. Nesta tarefa também está a ser desenvolvido um estudo integrado de modelos de EDE para o crescimento de organismos individuais (animais, plantas) em ambiente aleatório, incluindo instrumentos para aplicação prática. Estes estudos terão um impacto relevante na florestação e a criação de gado, actividades económicas muito importantes na região Alentejo. Participam principalmente nesta tarefa os colegas Carlos Braumann e Patrícia Filipe.

Tarefa 2: Estimação da abundância animal

Pretende-se o desenvolvimento e avaliação de modelos para estimar a abundância de populações animais e populações humanas com carácter evasivo e estimar parâmetros demográficos relacionados (taxas de mortalidade e sobrevivência e fluxos migratórios). Desenvolvem-se modelos combinados captura-recaptura e amostragem por distância sob enfoques frequencistas e bayesianos. Combinam-se também os modelos de transectos lineares e pontuais e avaliam-se os pressupostos tradicionais sob distribuições espaciais não homogéneas. O desenvolvimento destes modelos tem sido inspirado em aplicações como a estimação de populações de aves e de ungulados de montanha, e em populações humanas evasivas, como as que sofrem de doenças crónicas. As técnicas de captura-recaptura também foram propostas para estudar a relação entre probabilidades de extinção e probabilidades de detecção de espécies em dinâmica de comunidades. Participam principalmente nesta tarefa os colegas Russell Alpizar-Jara, Anabela Afonso e João Filipe Monteiro.

Tarefa 3: Métodos Multivariados Aplicados às Ciências da Vida

Nesta tarefa trabalha-se no desenvolvimento e aplicação de abordagens inferenciais a técnicas multivariadas que originalmente eram meramente descritivas (como por exemplo, extensões da metodologia STATIS e dos índices ambientais L2). De realçar o desenvolvimento de técnicas estatísticas sobre quase-normalidade na modelação de séries emparelhadas com estrutura generalizada comum, de inferência estatística para modelos mistos e de algoritmos para Análise de Regressão Conjunta com aplicações à teoria de selecção cultivar e melhoramento de plantas. Membros nesta tarefa também trabalham em técnicas de redução da dimensionalidade para critérios simples e múltiplos quase-óptimos de selecção de variáveis em Análise de Componentes Principais e em Modelos Lineares Generalizados. No âmbito destes trabalhos foi desenvolvido o componente de software *Subselect* para o *Projecto R*. Os métodos desenvolvidos nesta tarefa têm sido aplicados nas áreas da Biologia, Epidemiologia e Agricultura, entre outras. Participam principalmente nesta tarefa os colegas Manuela Oliveira, Dulce Gamito e Manuel Minhoto.

Tarefa 4: Processos Estocásticos em Ciências da Vida

Nesta tarefa inclui-se o estudo da estacionariedade fraca, de condições de existência e propriedades assintóticas em modelos de séries temporais de valores inteiros com parâmetros aleatórios, bem como testes de bondade do ajustamento e critérios de selecção em modelos do tipo DSINAR e RCINAR. Uma das aplicações foi a modelação do número de casos de mortes no distrito de Évora, considerando como covariáveis as temperaturas máximas e mínimas diárias. Outra área de actividade é a generalização de modelos para redes de telecomunicação de banda larga sem fios e para redes *ad hoc*, utilizando processos de Poisson não homogéneos. Além das aplicações na optimização de redes de telemóveis, há interesse nas aplicações à estimação de parâmetros e da distribuição espacial de populações animais. Participam principalmente nesta tarefa os colegas Dulce Gomes e Gonçalo Jacinto (em colaboração com membros envolvidos na tarefa 2).

Tarefa 5: Controlo de Qualidade e Análise de Sobrevivência em Ciências da Vida

Um dos objectivos desta tarefa consiste no desenvolvimento e combinação de métodos de amostragem em Controlo Estatístico da Qualidade e comparações das performances de novas metodologias com outras já existentes. A aplicação de modelos de captura-recaptura, tipicamente utilizados para estimar populações evasivas móveis, também é utilizada para estudos de fiabilidade em inspecção de software e, nesta tarefa, exploram-se ligações com o Controlo de Qualidade. Outro dos objectivos nesta tarefa é o estudo de modelos de sobrevivência espaciais multivariados e multi-estados com uma abordagem bayesiana na estimação de parâmetros. Estes modelos são motivados pelas aplicações na área da Saúde, como, por exemplo, a influência de factores climáticos (condições meteorológicas, poluição) no número de chamadas hospitalares devido a problemas respiratórios das crianças. Participam principalmente nesta tarefa os colegas Paulo Infante (em colaboração com membros envolvidos na tarefa 2) e Inês Dias.

Serviço à Comunidade Científica

A equipa tem sido muito activa quanto à participação em eventos nacionais e internacionais. Em particular, os membros da equipa têm vindo a participar muito activamente nos Congressos Anuais organizados pela Sociedade Portuguesa de Estatística (SPE). Em 2004 organizaram o XII Congresso Anual da SPE em Évora. Recentemente membros da equipa organizaram em Évora dois workshops internacionais em *Estimating Animal Abundance – WEAA* (7-9 de Julho de 2008) e *Dynamical Systems in Biology* (7 de Março de 2008). Os membros também têm participado em Comissões Organizadoras e/ou Científicas de vários congressos nacionais e internacionais. Por exemplo, nas Jornadas de Classificação e Análise de Dados (JOCLAD) e outros encontros nacionais, na 56ª Sessão do Instituto Internacional de Estatística (ISI 2007), na *Computational and Mathematical Population Dynamics* (CMPD) por duas vezes, no *Congreso Internacional do género Capra en Europa* e no *Symposium on Forecasting*, entre outros.

O Coordenador, Professor Carlos Braumann, é actualmente o Presidente da SPE e outros membros participam ou têm participado na Direcção e na Assembleia Geral da SPE. O Coordenador foi membro do Conselho Superior de Estatística e do CEIES (órgão consultivo da União Europeia para a estatística) e é actualmente membro da Direcção da *European Society for Mathematical and Theoretical Biology* e do *European Regional Committee* da *Bernoulli Society*.

Os membros da equipa são frequentemente chamados a participar como *referees* em revistas (*Bernoulli*, *Biometrics*, *J. Environmental and Ecological Statistics*, *J. of Agricultural, Biological and Environmental Statistics*, *J. Mathematical Biology*, *Mathematical Biosciences*, *International J. Mathematics and Mathematical Sciences*, *J. Theoretical Biology*, *J. Wildlife Management*, onde Russell Alpizar-Jara foi Editor Associado 2004-07, *J. Biodiversity and Conservation*, *J. Animal Ecology*, *European Journal of Operational Research*, *Interface-The Royal Society*, *Mathematics and Computers in Simulation*, *Computers & Mathematics with Applications*, *TEMA*, *Natural Resource*

Modeling, etc.) e em actas internacionais e nacionais, bem como reviewers da *Mathematical Reviews* e da *Zentralblatt Math*.

Colaboração Institucional

Membros da equipa mantêm colaboração com instituições nacionais e internacionais quer através do projecto em que se inserem no CIMA-U.E., quer por intermédio de outros projectos afins em que os membros participam e/ou coordenam. Por exemplo, o CIMA-U.E. mantém uma estreita interacção com o Departamento de Matemática da Universidade de Évora, quer na organização de actividades de divulgação e promoção da Estatística e da Matemática, quer através da organização de seminários conjuntos, quer na reestruturação e organização dos ensinos, particularmente, dando apoio à reestruturação dos mestrados em Modelação Estatística e Análise de Dados, mestrado em Matemática e Aplicações e programa de Doutoramento em Matemática, provendo também um espaço para o acolhimento de estudantes de pós-graduação, especialmente alunos de doutoramento e pós-doutoramento.

O CIMA-U.E. colabora como outros Centros da Universidade de Évora, como o Centro de Investigação em Ciências e Tecnologias da Saúde (CICT&S), o Centro de Geofísica (CGUE), o Instituto Ciências Agrárias Mediterrânicas (ICAM), o Centro de Estudos e Formação Avançada em Gestão e Economia (CEFAG), e o Centro Interdisciplinar de História, Culturas e Sociedades (CIDEHUS), nalguns casos, no âmbito de protocolos institucionalmente estabelecidos. Há também colaboração muito activa com centros da Universidade Nova de Lisboa (Faculdade de Ciências e Tecnologia) e da Universidade Técnica de Lisboa (Instituto Superior de Agronomia e Instituto Superior Técnico). Membros do projecto também têm mantido colaboração no âmbito da investigação e desenvolvimento tecnológico com o Instituto Nacional de Investigação Agrária e Pescas (INIAP-IPIMAR), o Instituto de Investigação da Floresta e Papel (RAIZ), o Grupo Soporcel-Portucel e o Instituto de Telecomunicações.

Internacionalmente, a equipa tem protocolos de colaboração com o Grupo de Biomatemática da Universidade de Campinas e com a Universidade Federal de Minas Gerais no Brasil, e com o Departamento de Estatística da Universidade de Costa Rica. Membros da equipa também têm mantido colaboração com as Universidades de Jaén, Extremadura, Santiago de Compostela, Politécnica de Valência e Politécnica de Madrid em Espanha, com a Universidade Paul Sabatier, Toulouse III, França, com as Universidades de *Brown*, *Thomas Jefferson*, *State University of New York at Stony Brook*, *North Carolina State* e com o *Patuxent Wildlife Research Center* nos Estados Unidos da América, com as Universidades de Zielona Gora e Poznan da Polónia, e com a Universidade de Joensuu da Finlândia.

O CIMA-U.E. é membro associado do Centro Internacional de Matemática (CIM) e mantém um Consórcio com várias unidades de Investigação Portuguesas (Minho, Algarve, Madeira e o antigo Instituto Superior de Educação de Cabo Verde) para partilhar a subscrição ao MathSciNet.

Produção Científica

A seguir apresentam-se alguns indicadores da produção do projecto SSALS do CIMA-U.E. nos últimos 4 anos (2003/07), que extraímos do Relatório para a FCT. Uma lista de publicações recentes e informação mais detalhada sobre o Centro está disponível no portal: <http://www.cima.uevora.pt/>.

Os membros do grupo têm participado em diversos projectos de investigação com financiamento exterior, quer coordenados por investigadores do grupo, quer por investigadores de outras unidades de investigação

Publicações nos últimos 4 anos	
Artigos em revistas internacionais com arbitragem	27
Artigos em revistas nacionais com arbitragem	1
Artigos em actas internacionais com arbitragem	21
Artigos em actas nacionais com arbitragem	36
Capítulos de livros internacionais com arbitragem	1
Capítulos de livros internacionais por convite	2
Capítulos de livros nacionais por convite	1
Livros (dois deles como editores)	3
Manuais didácticos	1
Total	93

Comunicações nos últimos 4 anos	
Comunicações convidadas em reuniões científicas internacionais (incluem-se 2 minicursos convidados)	12
Outras comunicações em reuniões científicas internacionais	45
Comunicações convidadas em reuniões científicas nacionais (incluem-se 2 minicursos convidados)	8
Outras comunicações em reuniões científicas nacionais	54
Seminários em instituições internacionais	9
Seminários em instituições nacionais	4
Total	132

Conclusão

Esperamos ter dado ao leitor uma panorâmica, necessariamente muito breve, da actividade no âmbito da Estatística desenvolvida no Centro de Investigação em Matemática e Aplicações da Universidade de Évora pelo Grupo SSALS (*Stochastic and Statistical Applications in Life Sciences*). Estamos convictos que o conhecimento mútuo do que se vem fazendo nesta área em Portugal ajudará a desenvolver a cooperação científica entre equipas de investigação com interesses comuns. Podemos ser visitados em <http://www.cima.uevora.pt/>.



CEMAT - Universidade Técnica de Lisboa

António Pacheco, apacheco@math.ist.utl.pt

*Centro de Matemática e Aplicações
Instituto Superior Técnico
Universidade Técnica de Lisboa*

O CEMAT e a Estatística no CEMAT

Neste documento faz-se um breve historial do CEMAT (Centro de Matemática e Aplicações), centro de investigação do IST (Instituto Superior Técnico – Universidade Técnica de Lisboa), e descrevem-se os seus objectivos e actividades principais, com relevo para os relacionados com a área de Estatística e Processos Estocásticos, a qual tem sido uma das suas áreas principais de investigação desde a sua formação.

Com o fim de não alongar demasiado a exposição, optou-se por efectuar uma apresentação generalista do Centro, mas que contém contudo uma grande quantidade de informação de interesse sobre o CEMAT, compilada de diversos relatórios do Centro produzidos no passado. A informação aqui detalhada pode e deve ser complementada com uma visita às páginas do CEMAT na *Internet*: <http://cemat.ist.utl.pt/>

1. Breve historial do CEMAT

O CEMAT é uma unidade de investigação financiada pela FCT (Fundação para a Ciência e a Tecnologia) desde a sua criação em Janeiro de 1996, ainda que o seu nome de baptismo tenha sido de Centro de Matemática Aplicada, o qual se manteve até 2001. Aquando da sua criação, o Centro possuía cerca de 60 membros, dos quais 24 eram doutorados, que se dividiam por quatro grupos de investigação:

- Grupo de Estatística, Processos Estocásticos e Aplicações
- Grupo de Análise Numérica e Aplicações em Mecânica de Fluidos
- Grupo de Teoria de Operadores, Equações Integrais e Aplicações
- Grupo de Física Matemática

Estes grupos eram liderados por João Branco, Adélia Sequeira, António Ferreira dos Santos e Roger Picken e contavam com 4, 6, 11 e 3 membros doutorados, respectivamente.

É importante realçar que os três primeiros grupos referidos, os quais despoletaram o processo de criação do Centro, resultaram de três linhas de investigação do CAPS (Centro de Análise e Processamento de Sinais do IST) que agrupavam investigadores da área de Matemática e existiam conjuntamente nesse centro desde 1994. A primeira Comissão Directiva do Centro incluía um membro de cada um destes três grupos, sendo a sua constituição a seguinte:

- António Ferreira dos Santos (Presidente)
- João Branco (Vice-Presidente)
- Carlos Alves (Vogal)

A importância do cargo desempenhado por João Branco, à altura Presidente da SPE, nessa Comissão espelha o papel relevante que ele desempenhou na criação do Centro.

Para além dos quatro grupos de investigação originais, o Centro integrou nos anos de 1997 a 2000 o Grupo de Lógica e Computação (liderado por Amílcar Sernadas), o qual proveio do ISR (Instituto de Sistemas e Robótica). Este grupo veio a promover a criação do CLC (Centro de Lógica e Computação do IST), tendo no início de 2001 os seus membros passado a integrar esse centro. Pese embora tenha sido bem acolhido pela FCT, o CLC veio a ser extinto em 2006, tendo a maioria dos seus membros passado a integrar o SQIG (Grupo de Segurança e Informação Quântica) do Instituto de Telecomunicações.

Para além da entrada e saída do Grupo de Lógica e Computação, outras mudanças na estrutura de grupos ocorreram no CEMAT. Em particular, o Grupo de Física Matemática, que foi sempre um grupo de dimensão relativamente pequena durante o período de permanência no Centro - tendo tido no máximo sete membros doutorados, decidiu no final de 2004 juntar-se a outro grupo da mesma área sediado no CAMGSD (Centro de Análise Matemática, Geometria e Sistemas Dinâmicos do IST).

Já anteriormente, mais precisamente no final de 2002, o Grupo de Teoria de Operadores, Equações Integrais e Aplicações dividiu-se em dois grupos: o Grupo de Teoria de Operadores, Álgebras de Banach e Aplicações (liderado por António Ferreira dos Santos) e o Grupo de Análise Harmónica, Teoria de Operadores e Aplicações (liderado por Stefan Samko), o primeiro destes grupos reunindo maioritariamente membros do IST e o segundo membros da Universidade do Algarve. Entretanto em 2006, estes grupos decidiram iniciar o caminho para criar um centro temático na área de Análise Funcional, dando deste modo origem ao actual CEAF (Centro de Análise Funcional e Aplicações do IST), o qual foi reconhecido pela FCT na última avaliação de centros de investigação que terminou em 2008.

Concorrentemente com a criação do CEAF e tendo em vista dar maior visibilidade às actividades de âmbito multidisciplinar em que o CEMAT tem estado envolvido, foi decidido criar o Grupo de Aplicações à Engenharia e a Biosistemas. Na sequência desta transformação, o Painel de Avaliação da Área de Matemática na última avaliação de centros de investigação efectuada pela FCT escreveu “CEMAT is now unified in the pursue of strengthening the applications of Numerical Analysis, Statistics and Stochastic Processes to Engineering and Biosystems.”

Em virtude das alterações ocorridas na estrutura do Centro ao longo do tempo, com entradas e saídas de grupos, ocorreram mudanças notórias no tamanho do CEMAT desde a sua criação. Partindo de 24 membros doutorados e um total de cerca de 60 membros no início de 1996, o Centro terá atingindo a sua dimensão máxima de 50 membros doutorados e um total cerca de 90 membros em 2004, imediatamente antes da saída do Grupo de Física Matemática para o CAMGSD. Uma dimensão próxima dessa terá sido atingida no início de 2001, antes da fundação do CLC. Já na altura da criação do CEAF, em 2006, o Centro possuía 43 membros doutorados e um total de 68 membros, tendo passado a ter 21 membros doutorados e um total de 35 membros após a criação do CEAF.

Se bem que a criação do CEAF tenha provocado a redução da dimensão do CEMAT para cerca de metade, o CEMAT gozou de um crescimento acelerado durante os últimos dois anos tendo neste curto espaço de tempo aumentado a sua dimensão em 40%, tendo passado para os actuais 49 membros, dos quais 33 são doutorados, possuindo 5 deles Agregação (Carlos Alves, Pedro Lima, António Pacheco, Ana Pires e Adélia Sequeira). Para este crescimento contribuíram o recrutamento de novos estudantes de pós-graduação e bolsiros de pós-doutoramento, a par de dois investigadores do Programa Ciência 2007, bem como a conclusão por membros do CEMAT de 9 teses de doutoramento nos últimos dois anos. As mudanças ocorridas durante os últimos três anos no CEMAT provocaram também um acentuado rejuvenescimento do Centro, tendo-o tornado um centro de investigação jovem e dinâmico, o qual conta apenas cinco membros com mais de cinquenta anos de idade.

Não queremos terminar esta secção sem deixar de salientar o facto de ao longo da sua existência o CEMAT ter dado um grande contributo para a produção científica nacional na área da matemática e suas aplicações, e ter dinamizado um grande número de actividades científicas na área. Em particular o trabalho científico dos membros do CEMAT conduziu: à publicação de mais de 400 artigos em revistas científicas internacionais; à escrita ou orientação de cerca de 50 dissertações de doutoramento e um número maior de dissertações de mestrado; à organização de um elevado número de encontros

científicos e outras acções de extensão universitária; a um grande número de acções de divulgação científica com ênfase em iniciativas do âmbito dos programas “Ciência Viva” e “O Livro da Natureza”; e ao lançamento e organização de várias séries de seminários regulares no IST, com relevo para os seguintes:

- Seminário de Análise Funcional e Aplicações
- Seminário de Física Matemática
- Seminário de Lógica e Computação
- Seminário de Matemática Aplicada e Análise Numérica
- Seminário de Probabilidades e Estatística

Para finalizar a secção, apresenta-se a lista dos presidentes do CEMAT até à actualidade:

- António Ferreira dos Santos (1996-1998)
- Cristina Sernadas (1999-2000)
- Adélia Sequeira (2001-2006)
- António Pacheco (desde 2007)

2. Composição, natureza, organização e principais actividades do CEMAT

O CEMAT é um centro de investigação em Matemática cujo objectivo principal é o desenvolvimento de investigação científica de elevada qualidade em matemática aplicada, com relevo para as áreas de Estatística, Processos Estocásticos, Análise Numérica e Matemática Computacional, promovendo concomitantemente o enriquecimento científico dos seus membros e colaboradores.

No seu conjunto, o CEMAT integra actualmente 49 membros, dos quais 33 são doutorados, para além de vários colaboradores. Do conjunto de 33 membros doutorados: 26 são professores universitários em exercício de funções, 2 são investigadores do Programa Ciência 2007, 4 são bolseiros de pós-doutoramento e 1 é consultor em empresa privada. Os 26 professores universitários do CEMAT em exercício de funções estão repartidos por diferentes instituições universitárias do seguinte modo: Instituto Superior Técnico (15), Instituto Politécnico de Leiria (4); Universidade do Algarve (2); Universidade de Trás-os-Montes e Alto Douro (2), Instituto Superior de Economia e Gestão (1), Instituto Superior de Engenharia de Lisboa (1) e Universidade do Porto (1).

Em resultado da sua composição, o CEMAT afirma-se como um centro de investigação de âmbito nacional mas maioritariamente associado ao IST. A nível de dimensão, constitui no panorama nacional um centro de investigação de média dimensão, tanto a nível geral como na área de Matemática. Ao longo da sua existência o CEMAT tem exibido elevados índices de produtividade científica em todas as áreas onde tem desenvolvido investigação, tendo recebido sempre a classificação de ‘Muito Bom’ nas avaliações de centros de investigação realizadas pela FCT.

A investigação realizada no CEMAT é actualmente enquadrada em três grupos de investigação:

- Grupo de Estatística e Processos Estocásticos
- Grupo de Análise Numérica e Análise Aplicada
- Grupo de Aplicações à Engenharia e a Biosistemas

liderados por António Pacheco, Carlos Alves e Adélia Sequeira, respectivamente. Os dois primeiros grupos de investigação decorrem directamente de grupos existentes quando da criação do Centro e agregam a investigação centrada mais exclusivamente em aspectos de modelação e análise matemática. Já o terceiro grupo de investigação visa sobretudo realçar o enfoque aplicado da investigação realizada no CEMAT, que incentiva fortemente o envolvimento dos seus membros em equipas/acções multidisciplinares em que a modelação matemática desempenha um papel importante. Convém realçar que a maioria dos membros deste último grupo são também membros de um dos outros dois grupos de investigação.

Recentemente, a investigação realizada no CEMAT tem estado centrada principalmente nos seguintes tópicos:

- Análise Aplicada
- Análise e Simulação Numérica em Mecânica dos Fluidos
- Análise Multivariada
- Análise Numérica de Equações Diferenciais e Integrais Ordinárias
- Computação Paralela
- Controlo de Qualidade
- Estatística Computacional
- Estatística Robusta
- Genómica Estatística
- Métodos *Meshless* e Aplicações
- Modelação Matemática e Computacional em Medicina
- Modelação Matemática em Hemodinâmica
- Modelação Matemática e Numérica de Escoamentos Geofísicos
- Modelos Estocásticos em Telecomunicações
- Ordenação Estocástica
- Problemas Inversos e Aplicações
- Redes de Filas de Espera

A estrutura organizativa do CEMAT está orientada a centrar a actividade dos seus membros na investigação, libertando-os ao máximo de actividades de gestão científica, sendo materializada numa estrutura de gestão leve assente nos seguintes órgãos:

- Comissão Executiva
- Comissão Coordenadora
- Conselho Científico

O CEMAT é administrado pela Comissão Executiva, constituída actualmente por:

- António Pacheco (Presidente)
- Adélia Sequeira (Vice-Presidente)
- Ana Pires (Vogal)

O Presidente do CEMAT coordena a definição do projecto de investigação do Centro e respectiva implementação; sendo eleito para um mandato de dois anos, é ele que designa os restantes membros da Comissão Executiva que o apoiam na sua missão. O Conselho Científico é constituído pelos membros doutorados do CEMAT e elege o Presidente, ratifica a Comissão Executiva e aprova alterações aos estatutos e à criação e dissolução de grupos de investigação. A Comissão Coordenadora (do Conselho Científico), constituída pelos membros da Comissão Executiva e pelos coordenadores dos grupos de investigação, aprova os relatórios anuais de actividades e colabora com o Presidente na definição das políticas de investigação do CEMAT.

O CEMAT possui ainda uma Comissão de Acompanhamento, a qual emite pareceres e recomendações sobre as actividades e políticas de investigação do CEMAT; integram actualmente esta comissão os seguintes cinco professores:

- Ivan Graham (Universidade de Bath, Reino Unido)
- Guy Latouche (Universidade Livre de Bruxelas, Bélgica)
- Alfio Quarteroni (EPFL e Politécnico de Turim, Suíça e Itália)
- Peter Rousseuw (Universidade de Antuérpia, Bélgica)
- J. George Shanthikumar (Universidade da Califórnia em Berkeley, EUA)

A par da realização de investigação de qualidade na área de matemática aplicada e sua publicação em literatura especializada, com relevo para revistas internacionais, são também primordiais para o CEMAT as seguintes actividades:

- Formação pós-graduada, com ênfase na doutoral e pós-doutoral
- Envolvimento em redes e projectos científicos interdisciplinares e multidisciplinares
- Organização de encontros científicos e acções de divulgação científica
- Organização de séries de seminários

O CEMAT estimula fortemente a formação pós-graduada, estando membros seus envolvidos de modo muito visível na leccionação de unidades curriculares e orientação de dissertações em cursos de mestrado e doutoramento nas áreas de investigação do Centro, como detalhado nos parágrafos seguintes. Por outro lado, possui um activo programa pós-doutoral, o qual dá primazia ao envolvimento dos seus membros em projectos de investigação em que o CEMAT participa. Este programa envolveu 14 doutores desde 2004 e inclui presentemente quatro doutores.

A nível de cursos de mestrado o maior envolvimento do CEMAT ocorre actualmente no Mestrado em Matemática e Aplicações do IST, se bem que esteja também envolvido noutros cursos do IST, como sejam os mestrados em Engenharia Biomédica e em Engenharia de Redes. No passado recente, o CEMAT teve um papel primordial no Mestrado em Estatística do IST, curso pré-Bolonha que já não se encontra em funcionamento. Desde 2007, foram defendidas 23 dissertações de mestrado orientadas por membros do CEMAT.

O CEMAT possui um programa doutoral muito estimulante, em que promove o envolvimento activo dos seus doutorandos em actividades de investigação a partir das fases iniciais dos seus estudos de doutoramento, bem como a participação, com apoio financeiro por parte do CEMAT, em conferências e cursos científicos de interesse para o respectivo programa de trabalho. Presentemente, os participantes no programa doutoral do CEMAT são integrados primordialmente nos seguintes dois cursos do IST:

- Doutoramento em Estatística e Processos Estocásticos
- Doutoramento em Matemática

a par dos seguintes quatro cursos em associação com outras universidades:

- Doutoramento em Engenharia e Ciências Computacionais (UT Austin-Portugal)
- Doutoramento em Matemática Aplicada do ICTI (CMU-Portugal)
- Doutoramento em Matemática Computacional e Estocástica (IST-EPFL)
- Doutoramento em Matemática (UT Austin-Portugal)

Desde 2007, foram concluídas 9 dissertações de doutoramento no âmbito do CEMAT.

O CEMAT promove muito activamente a realização de projectos de investigação em colaboração com outras instituições, incluindo unidades de investigação não pertencentes à área da Matemática, como forma de fortalecer no CEMAT a área de aplicações. Um bom exemplo disto é dado pela colaboração do CEMAT com o Instituto de Telecomunicações em três projectos financiados pela FCT na área de modelação e análise estatística de tráfego de telecomunicações: SCALE (Statistical Characterization of Telecommunications Traffic - POSI/34826/CPS/2000); PERMNET (Traffic Modeling and Performance Evaluation of Multiservice Networks - POSI/42069/CPS/2001); e MOMSAIT (Internet Traffic Measurements, Modelling and Statistical Analysis – POSC/EIA/60061/2004). Da investigação levada a cabo nestes projectos resultou a publicação de um grande número de artigos científicos dispersos por vários periódicos internacionais, incluindo nomeadamente: *Advances in Applied Probability*; *Annals of Operations Research*; *Computer Networks*; *Computers and Mathematics with Applications*; *IEICE Transactions on Communications*; *Probability in the Engineering and Informational Sciences*; *Queueing Systems*; *Performance Evaluation*; *Stochastic Models e Telecommunication Systems*.

O CEMAT estimula também o envolvimento dos seus membros em redes científicas internacionais; a este respeito, equipas do CEMAT encontram-se ou estiveram recentemente envolvidas, em particular, nas três seguintes redes:

- **EUCANET** (Eucalyptus Genomics Research Network for Improved Wood Properties and Adaptation to Drought) - rede envolvendo várias instituições de Portugal, Espanha e França, financiada por INRA (Institut National de la Recherche Agronomique, França), FCT (Fundação para a Ciência e a Tecnologia, Portugal) e MEPSYD (Ministerio de Educación, Política Social y Deporte, Espanha). Esta rede visa, com enquadramento a nível da Genómica Estatística, a identificação de um número restrito de genes candidatos como susceptíveis de melhorarem as características de produtividade de madeira de eucalipto, uma das principais espécies florestais no Sul da Europa, sob condições de seca. A participação do CEMAT nesta rede é liderada por **Ana Pires**.
- **Euro-FGI** (Design and Engineering of the Future Generation Internet) – Rede de Excelência da União Europeia envolvendo 56 instituições de 19 países europeus, orientada ao desenvolvimento de arquiteturas de rede e metodologias quantitativas de análise de tráfego inovadoras adaptadas às exigências da Internet do futuro. A participação do CEMAT nesta rede foi liderada por **António Pacheco**.
- **HaeModel** (Mathematical and Numerical Modelling in Haemodynamics, Research Training Network) - rede envolvendo CEMAT, EPFL (Suíça), Imperial College (Reino Unido), INRIA (França), Politecnico di Milano (Itália) e a Universidade Técnica de Graz (Áustria), financiada por EU-RTN (European Union Research Training Network) e OFES (Swiss National Bureau of Science and Education). Esta rede esteve orientada ao desenvolvimento de ferramentas para reconstrução geométrica a partir de dados médicos e ao desenvolvimento de modelos matemáticos de interacção de transporte bioquímico e mecânica bioquímica em vasos sanguíneos. A participação do CEMAT nesta rede foi liderada por **Adélia Sequeira**.

Os membros do CEMAT têm um grande envolvimento a nível de organização e colaboração na realização de encontros científicos e acções de divulgação científica. Dois exemplos relevantes de importantes conferências internacionais com envolvimento do CEMAT são:

- **IWANASP08** (Third International Workshop on Analysis and Numerical Approximation of Singular Problems, Ericeira, 10-12 Setembro, 2008). url: <http://www.math.ist.utl.pt/~tdiogo/IWANASP08/> Comissão Organizadora: **Teresa Diogo, Pedro Lima, Magda Rebelo e Filomena Teodoro**.
- **ICORS 2006** (International Conference on Robust Statistics 2006, Lisboa, 16-21 Julho, 2006); url: <http://www.math.ist.utl.pt/icors2006/> Comissão Organizadora: **Conceição Amado, Rosário Oliveira, Carla Pereira, Ana Pires, Isabel Rodrigues, Peter Rousseeuw e Manuela Souto de Miranda**.

Já a nível nacional, são particularmente relevantes, para além doutras iniciativas, as participações de membros do CEMAT na organização do VI Congresso Anual da SPE, realizado em Tomar em 1998, por intermédio de **António Pacheco, Ana Pires e Daniel Paulino**, e do XVI Congresso Anual da SPE, realizado em Vila Real em 2008, por intermédio de **Fátima Ferreira**. Um exemplo recente de contribuição do CEMAT a nível de divulgação científica foi o envolvimento de **Cláudia Nunes e Adélia Sequeira** na EIM09 (Escola de Inverno de Matemática, Instituto Superior Técnico, Lisboa, 9-14 Fevereiro, 2009).

Em continuidade com a tradição do CEMAT de dinamismo na organização de seminários de investigação, envolvendo a participação regular de investigadores estrangeiros de renome, o Centro organiza actualmente as seguintes três séries de seminários do CEMAT:

- Seminário de Probabilidades e Estatística
- Seminário de Matemática Aplicada e Análise Numérica
- Seminário Aberto

Para além disso, o CEMAT colabora na organização de:

- Colóquio do Departamento de Matemática do IST
- Seminário de Probabilidades e Estatística da UTL
- Seminários e *workshops* do CIM (Centro Internacional de Matemática)

3. Breve historial da evolução do Grupo de Estatística e Processos Estocásticos

João Branco coordenou a formação em 1993 do Grupo de Estatística, Processos Estocásticos e Aplicações do IST. Este grupo, que passou a integrar o CAPS a partir de 1994, incluía um total de 16 membros do IST: três doutorados (João Amaral, João Branco e Daniel Paulino), três mestres (António Pacheco, Ana Pires e Giovani Silva), nove licenciados (Conceição Amado, António Anjos, Francisco Conceição, Manuel Morais, Cláudia Nunes, Rosário Oliveira, Isabel Rodrigues, Fernando Sequeira e Paulo Soares) e um não-licenciado (Duarte Gonçalves).

No que se segue, daremos uma breve ideia da evolução desde o grupo referido até ao actual Grupo de Estatística e Processos Estocásticos, referindo-nos sempre ao grupo em questão como GEPE. Para registo, começa-se por referir que à data da criação formal do CEMAT, em Janeiro de 1996, a constituição do GEPE tinha-se já alterado em relação ao descrito no parágrafo anterior pela entrada de dois não-licenciados, Rui Paulo e Susana Segão, e pelas saídas de António Anjos, Francisco Conceição, Fernando Sequeira e Duarte Gonçalves, contando um total de 14 membros.

Como deixa antever a descrição apresentada no início desta secção, quando da sua criação o GEPE era um grupo muito jovem e relativamente pouco qualificado. Este facto justificou plenamente que a actividade inicial do GEPE tenha sido orientada sobretudo à formação dos seus membros não-doutorados. Essa actividade foi materializada no doutoramento dos seus três mestres iniciais (António Pacheco, em 1994; Ana Pires, em 1996; e Giovani Silva, em 2001) e na conclusão por Conceição Amado, Manuel Morais, Cláudia Nunes, Rosário Oliveira, Isabel Rodrigues e Paulo Soares: numa primeira fase de mestrados (Conceição Amado, em 1994; Manuel Morais, Cláudia Nunes e Rosário Oliveira, em 1995; Paulo Soares, em 1996; e Isabel Rodrigues, em 1997) e numa segunda fase de doutoramentos (Manuel Morais e Rosário Oliveira, em 2002; Cláudia Nunes e Isabel Rodrigues, em 2003; e Conceição Amado e Paulo Soares, em 2004).

O esforço de formação de membros não-doutorados do GEPE foi estendido a outros membros integrados no CEMAT a partir de 1996 através da conclusão dos respectivos mestrados (Susana Segão, em 2000; Patrícia Ferreira e Nuno Sepúlveda, em 2006; Cláudia Pascoal, em 2007; Sílvia Nobre, em 2008; e Filipa Encarnação, em 2009) e doutoramentos (Nelson Antunes, em 2001; Rui Paulo, em 2002; Carla Pereira, em 2004; e Fátima Ferreira e Helena Ribeiro, em 2007). Convém ainda realçar que, numa fase recente, três membros do GEPE foram aprovados em provas de agregação: Daniel Paulino, em 2005; António Pacheco, em 2006; e Ana Pires, em 2008. O primeiro destes obteve a agregação em Estatística e Investigação Operacional, pela Faculdade de Ciências da Universidade de Lisboa, e os dois restantes em Matemática, pela Universidade Técnica de Lisboa.

Com excepção de uma tese de mestrado (Susana Segão) e três teses de doutoramento (António Pacheco, Rui Paulo e Giovani Silva), os mestrados e doutoramentos até agora mencionados foram orientados por membros do CEMAT, em acordo com a informação apresentada nos próximos dois parágrafos sobre orientações de dissertações de mestrado e doutoramento. Refira-se que António Pacheco e Rui Paulo obtiveram os seus doutoramentos nos Estados Unidos (nas universidades de *Cornell* e *Duke*, respectivamente) e Susana Segão e Giovani Silva foram orientados por Antónia Amaral-Turkman, da Faculdade de Ciências da Universidade de Lisboa, nos seus mestrado e doutoramento, respectivamente.

No seu conjunto, os membros doutorados do GEPE orientaram desde 1994 mais de trinta dissertações de mestrado e as doze dissertações de doutoramento discriminadas na Tabela 1. Estes números são especialmente relevantes tendo em conta que, por exemplo, ao longo da existência do GEPE nem sempre a área científica de Probabilidades de Estatística existiu como área de especialização de mestrado no IST.

Como se constata da Tabela 1, na formação de novos doutores pelo GEPE tiveram um papel muito especial três professores: João Branco, António Pacheco e Ana Pires. Já na orientação de dissertações de mestrado contribuíram quase todos os doutores do GEPE, como está patente na seguinte lista de nomes de mestres cujas dissertações de mestrado foram orientadas por membros do GEPE, estes últimos apresentados entre parêntesis: Alexandra Mayer da Silva (Rosário Oliveira), Ana Mota (Ana Pires), António Papulim (João Branco), Bartolomeu Rabaçal (Isabel Rodrigues), Bruno Silva (Cláudia

Nunes), Carla Miranda (Ana Pires), Carla Patrícia Santos (Conceição Amado), Catarina Rego (Conceição Amado), Catarina Soares (Ana Pires), Cláudia Antunes (Manuel Morais & António Pacheco), Cláudia Nunes (João Amaral), Cláudia Pascoal (Rosário Oliveira & António Pacheco), Delfina Barbosa (Conceição Amado), Filipa Encarnação (Rosário Oliveira & Luzia Gonçalves), Gonçalo Jacinto (António Pacheco & Nelson Antunes), Helena Pereira (Conceição Amado), Isabel Rodrigues (Daniel Paulino), Luís Sertório (Rosário Oliveira & Rui Valadas), Maria João Polidoro (João Branco), Maria José Castro (João Branco), Maria Raquel Neto (João Branco), Nuno Sepúlveda (Daniel Paulino & Carlos Penha-Gonçalves), Patrícia Ferreira (João Branco), Paulo Canas (João Branco), Paulo Soares (Daniel Paulino), Regina Soares (Ana Pires), Rosário Oliveira (João Branco), Rui Branco (Ana Pires), Sérgio Valente (João Branco), Sílvia Nobre (Cláudia Nunes), Sofia Pinto (Conceição Amado), Sónia N. Dória (Daniel Paulino), Tânia Silva (Cláudia Nunes & António Pacheco) e Teresa Riso (Ana Pires).

Merece especial destaque o facto de, com excepção de Manuela Souto (de Miranda) que permaneceu sempre como membro do Centro de Matemática e Aplicações da Universidade de Aveiro, todos os novos doutores orientados por membros do GEPE, listados na Tabela 1, terem ingressado no CEMAT antes do doutoramento e aí terem permanecido após a conclusão do respectivo doutoramento. Este facto é resultado das boas políticas de gestão de recursos humanos e promoção das actividades de investigação, inclusivas a membros não-doutorados, seguidas no CEMAT.

Por oposição ao ingresso no GEPE de novos doutores formados no CEMAT, registou-se a saída de cinco dos seus membros doutorados para outros centros de investigação a fim de seguirem projectos científicos diversos adequados aos seus próprios interesses de investigação: Rui Paulo para o CEMAPRE, em 2006, após o seu ingresso como professor auxiliar do ISEG; Daniel Paulino, Giovani Silva e Paulo Soares para o CEAUL, em 2007; e João Costa e Silva (que ingressou no CEMAT como bolsheiro de pós-doutoramento em 2007) para o Centro de Estudos Florestais do ISA, em 2009. Por outro lado, o GEPE acolheu em 2008, como investigador do Programa Ciência 2007, Vladas Pipiras que é professor associado na Universidade da Carolina do Norte em Chapel Hill.

Tabela 1: Lista de dissertações de doutoramento orientadas por membros do GEPE

Ano	Orientador	Orientando	Título da Dissertação
1995	João Branco	Manuela Souto	Estimação Robusta com Variáveis Instrumentais em Modelos com Erros-nas-Variáveis
1996	João Branco	Ana Pires	Análise Discriminante: Novos Métodos Robustos de Estimação
2001	António Pacheco	Nelson Antunes	Modeling and Analysis of Wireless Networks
2002	António Pacheco	Manuel Morais	Stochastic Ordering in the Performance Analysis of Quality Control Schemes
2002	João Branco	Rosário Oliveira	Métodos Robustos em Análise de Correlações Canónicas
2003	António Pacheco	Cláudia Nunes	Markov Modulated Chains
2003	Ana Pires	Isabel Rodrigues	Métodos Robustos em Análise de Componentes Principais Comuns
2004	Ana Pires	Conceição Amado	Bootstrap Robusto com base na Função de Influência
2004	Ana Pires	Carla Pereira	Reconhecimento de Padrões: Classificação Supervisionada com Rejeição de Observações
2004	Daniel Paulino	Paulo Soares	Análise Bayesiana de Dados Deficientemente Categorizados
2007	António Pacheco	Fátima Ferreira	Embedding, Uniformization and Stochastic Ordering in the Analysis of Level-Crossing Times and $GI^X/M(n)/c$ Systems
2007	António Pacheco	Helena Ribeiro	Customer Loss Probabilities and other Performance Measures of Regular and Oscillating Systems

Em resultado da formação de novos doutores no seio do GEPE e das entradas e saídas de membros doutorados ocorridas durante a sua existência, integram actualmente o GEPE os seguintes 14 doutores:

- Ana Pires (IST)
- António Pacheco (IST)
- Carla Pereira (UPorto)
- Cláudia Nunes (IST)
- Conceição Amado (IST)
- Fátima Ferreira (UTAD)
- Helena Ribeiro (IPLeiria)
- Isabel Rodrigues (IST)
- João Amaral (IST)
- João Branco (IST)
- Manuel Morais (IST)
- Nelson Antunes (UAlgarve)
- Rosário Oliveira (IST)
- Vladas Pipiras (IST e UNC-Chapel Hill)

O GEPE inclui ainda os seguintes sete mestres: Cláudia Pascoal, Elena Almaraz-Luengo, Elisabete Fernandes, Filipa Encarnação, Patrícia Ferreira, Sílvia Nobre e Vanda Lourenço. Em face da composição do GEPE, pode afirmar-se que o CEMAT possui actualmente um bom grupo de investigação em Estatística e Processos Estocásticos.

4. Actividade e resultados do Grupo de Estatística e Processos Estocásticos

A actividade e resultados do GEPE estão alinhados com a do CEMAT em geral, a qual foi relatada nas duas primeiras secções do documento; já a actividade de formação do GEPE foi amplamente detalhada na Secção 3. Assim, nesta secção limitar-nos-emos a complementar a informação das secções anteriores no que ao GEPE diz respeito.

Ao longo da sua existência, a actividade de investigação GEPE incidiu sobretudo sobre os seguintes tópicos:

- Análise de Sobrevivência
- Análise Multivariada
- Controlo de Qualidade
- Estatística Computacional
- Estatística Robusta
- Inferência Bayesiana
- Genómica Estatística
- Modelos Estocásticos de Computação
- Modelos Estocásticos em Telecomunicações
- Ordenação Estocástica
- Redes de Filas de Espera

Estes tópicos continuam a ser alvo de grande atenção actualmente, com excepção de Análise de Sobrevivência, Inferência Bayesiana e Modelos Estocásticos de Computação. Dois dos tópicos apresentados, Genómica Estatística e Modelos Estocásticos em Telecomunicações, estão ligados essencialmente às aplicações e são desenvolvidos em colaboração estreita com instituições ligadas às áreas biológica e de telecomunicações, nomeadamente: o Instituto Gulbenkian de Ciência e o Instituto de Investigação da Floresta e Papel (RAIZ) no referente à Genómica Estatística; e o Instituto de Telecomunicações no respeitante a Modelos Estocásticos em Telecomunicações.

A actividade de publicação de artigos científicos nas áreas referidas por parte de membros do GEPE está dispersa por mais de duzentos trabalhos, repartidas por publicações nacionais e internacionais. Começando por livros, os membros do GEPE foram autores ou co-autores de oito livros, incluindo três livros de mini-cursos de congressos anuais da SPE:

- M. A. Amaral Turkman e **G. Silva** (2000). *Modelos Lineares Generalizados – da teoria à prática*. Edições SPE, Lisboa.
- **C. D. Paulino**, M. A. Amaral Turkman e B. Murteira (2003). *Estatística Bayesiana*. Fundação Calouste Gulbenkian, Lisboa.
- **J. A. Branco** (2004). *Uma Introdução à Análise de Clusters*. Edições SPE, Lisboa.
- **C. D. Paulino** e **J. Branco** (2005). *Exercícios de Probabilidade e Estatística*. Escolar Editora, Lisboa.
- **C. D. Paulino** e J. M. Singer (2006). *Análise de Dados Categorizados*. Edgard Blucher, São Paulo.
- **G. L. Silva** e C. B. Dean (2006). *Introdução à Análise de Modelos Espaço-temporais para Taxas, Proporções e Processos de Multi-estados*. XVII SINAPE - Associação Brasileira de Estatística, São Paulo.
- **A. M. Pires** e **J. A. Branco** (2007). *Introdução aos Métodos Estatísticos Robustos*. Edições SPE, Lisboa.
- **A. Pacheco**, L. C. Tang e N. U. Prabhu (2008). *Markov-Modulated Processes and Semiregenerative Phenomena*. World Scientific, Singapura.

A nível de artigos com publicação nacional, a esmagadora maioria dos trabalhos de membros do GEPE foram publicados nas actas dos congressos da SPE, num total de cerca de 60 artigos, tendo membros do GEPE sido também editores das actas dos I, III e VI congressos anuais da SPE:

- *A Estatística e o Futuro e o Futuro da Estatística* - Actas do I Congresso Anual da SPE. D. Pestana, A. Turkman, **J. Branco**, L. Duarte e **A. Pires** (Eds.). Edições Salamandra, Lisboa, 1994.
- *Bom Senso e Sensibilidade – Traves Mestras da Estatística* - Actas do III Congresso Anual da SPE. **J. Branco**, P. Gomes e J. Prata (Eds.). Edições Salamandra, Lisboa, 1996.
- *Afirmar a Estatística – Um Desafio para o Século XXI* - Actas do VI Congresso Anual da SPE. **D. Paulino**, **A. Pacheco**, **A. Pires** e Ferreira da Cunha (Eds.). Edições SPE, Lisboa, 1999.

A nível internacional, a maioria das publicações de membros do GEPE são artigos em revista, tendo no seu conjunto sido autores ou co-autores de cerca de uma centena de artigos publicados em cerca de sessenta periódicos internacionais com recensão de artigos.

A maior parte das revistas internacionais em que os membros do GEPE têm publicado são da área de Probabilidades e Estatística e incluem, em particular, as seguintes: *Advances in Applied Probability*; *Annals of Applied Probability*; *Annals of Probability*; *Annals of Statistics*; *Biometrics*; *Biometrika*; *Chemometrics and Intelligent Laboratory Systems*; *Computational Statistics*; *Computational Statistics and Data Analysis*; *Economic Quality Control*; *Environmetrics*; *J. of Applied Probability*; *J. of Applied Statistics*; *J. of Classification*; *J. of Multivariate Analysis*; *J. of Statistical Computation and Simulation*; *Methodology and Computing in Applied Probability*; *Metrika*; *Probability in the Engineering and Informational Sciences*; *REVSTAT*; *Sequential Analysis*; *Statistics & Decisions*; *Statistics & Probability Letters*; *Stochastic Models* e *Stochastic Processes and their Applications*.

Por outro lado, mais de uma dezena de artigos de membros do GEPE em áreas de interface de Estatística e Processos Estocásticos com as áreas da Ciência da Computação, outras áreas de Matemática ou Investigação Operacional foram publicados em revistas dessas áreas, incluindo as seguintes: *Annals of Mathematics and Artificial Intelligence*; *Annals of Operations Research*; *Computers and Mathematics with Applications*; *Journal of Fourier Analysis and Applications*; *Queueing Systems* e *Theoretical Computer Science*. Já mais de duas dezenas de artigos de membros do GEPE em áreas de aplicação de Estatística e Processos Estocásticos, com relevo para as áreas biológicas e de telecomunicações, foram publicados em revistas das respectivas áreas de aplicação, incluindo as seguintes: *Computer Methods and Programs in Biomedicine*; *Computer Networks*; *Ecotoxicology and Environmental Safety*; *Forest Ecology and Management*; *IEICE Transactions on Communications*; *International Journal Communications in Dependability and Quality Management*; *Pathobiology*; *Pattern Recognition Letters*; *Performance Evaluation*; *Statistics in Medicine*;

O caminho trilhado pelo GEPE tem produzido bons resultados e conduzido ao seu reconhecimento pela comunidade estatística tanto a nível nacional como internacional. Um texto que ilustra bem o reconhecimento mencionado é a parte final do relatório produzido há poucos meses pelo Prof. Peter Rousseeuw, na qualidade de membro da Comissão de Acompanhamento do CEMAT, no tocante à investigação do GEPE em Estatística Robusta e ao CEMAT em geral, transcrita a seguir:

Over the years, the CEMAT group has worked on making robust statistics possible in Principal Component Analysis, Canonical Correlation Analysis, Discriminant Analysis, Multiple Regression with correlated errors, and several other multivariate or otherwise structured frameworks. In each case their work spans the full gamut from theoretical results (influence function, asymptotic, breakdown value) over algorithms and simulations to actual applications. The results were published in excellent journals like Biometrika and Journal of Multivariate Analysis, thereby achieving ample visibility and recognition. Some of this work was done in collaboration with other statisticians of high standing in the UK, Argentina, and Austria. One measure of how well the field of robust statistics is progressing is the fact that many techniques have been included in the standard version of major statistical packages like SAS, S-Plus, and R, thanks in part to the CEMAT group.

To conclude, in my view CEMAT is an active research institute which does important and interesting work, is well organized and integrated in the international community of applied mathematics, and is able to attract good students and post-doc researchers. It is still growing and improving over time and deserves encouragement and support for its further development.

O reconhecimento de que o CEMAT é alvo constitui um forte estímulo para dar continuidade, inovando, ao trabalho que tem vindo a desenvolver.



CEAUL - Universidade de Lisboa

Maria Isabel Fraga Alves (coordenadora), isabel.alves@fc.ul.pt

*Centro de Estatística e Aplicações da Universidade de Lisboa
Departamento de Estatística e Investigação Operacional
Faculdade de Ciências, Universidade de Lisboa*

Quem fomos, somos e para onde vamos

1. Introdução

O Centro de Estatística e Aplicações, sendo o maior e dos mais activos grupos de investigação nesta área, está bem posicionado para contribuir ao solicitado em aspectos variados das Ciências Estatísticas. De acordo com o plano de acção estabelecido por esta Unidade de Investigação para os próximos anos, o principal objectivo é o de contribuir para o avanço e uso adequado da Estatística em Portugal e pensa-se que isto pode ser conseguido através de várias actividades interligadas.

O CEAUL foi fundado em 1975 pelo Prof. J. Tiago de Oliveira, podendo dizer-se que foi o berço da investigação em Probabilidades e Estatística no nosso País. A grande maioria dos actuais Mestres e Doutorados nestas áreas, são (ou foram) elementos deste Centro, ou foram orientados por elementos que a ele pertencem. O CEAUL albergou no seu início outras áreas, tais como Ciência da Computação e Investigação Operacional. Devido ao grande desenvolvimento verificado nestas áreas durante a década de 80, deu-se a separação inevitável a área da Ciência da Computação criou autonomia em 1985; em 1993 deu-se a separação da área da Investigação Operacional, elementos da linha de acção em Optimização e Aplicações, já que o número dos seus elementos e a actividade científica por eles desenvolvida justificava plenamente a existência de uma Unidade independente; O CEAUL é uma unidade I&D junto do DEIO/FCUL, sendo esta a Instituição de Acolhimento, disponibilizando instalações e infra-estruturas necessárias à prossecução das suas actividades, gestão orçamental a cargo da FFCUL, facultando-lhe a colaboração de um técnico de secretaria vinculado à Unidade de Acolhimento FCUL, de forma a cumprir com o disposto no artigo 2 do Regulamento do Programa de Financiamento Plurianual de I&D – 2007. O financiamento do CEAUL é assegurado no quadro de PFP(I&D)FCT.

Past Coordenadores do CEAUL	
1975-1981	<i>José Tiago da Fonseca Oliveira</i>
1981-1987	<i>Dinis Duarte Ferreira Pestana</i>
1987-1992	<i>José Manuel Pinto Paixão</i>
1992-1999	<i>Kamil Feridun Turkman</i>
1999-2006	<i>Maria Ivette Leal de Carvalho Gomes</i>
2006-	<i>Maria Isabel Fraga Alves</i>



O CEAUL é membro da CNM-Comissão Nacional de Matemática/International Mathematical Union (IMU) – Portugal.



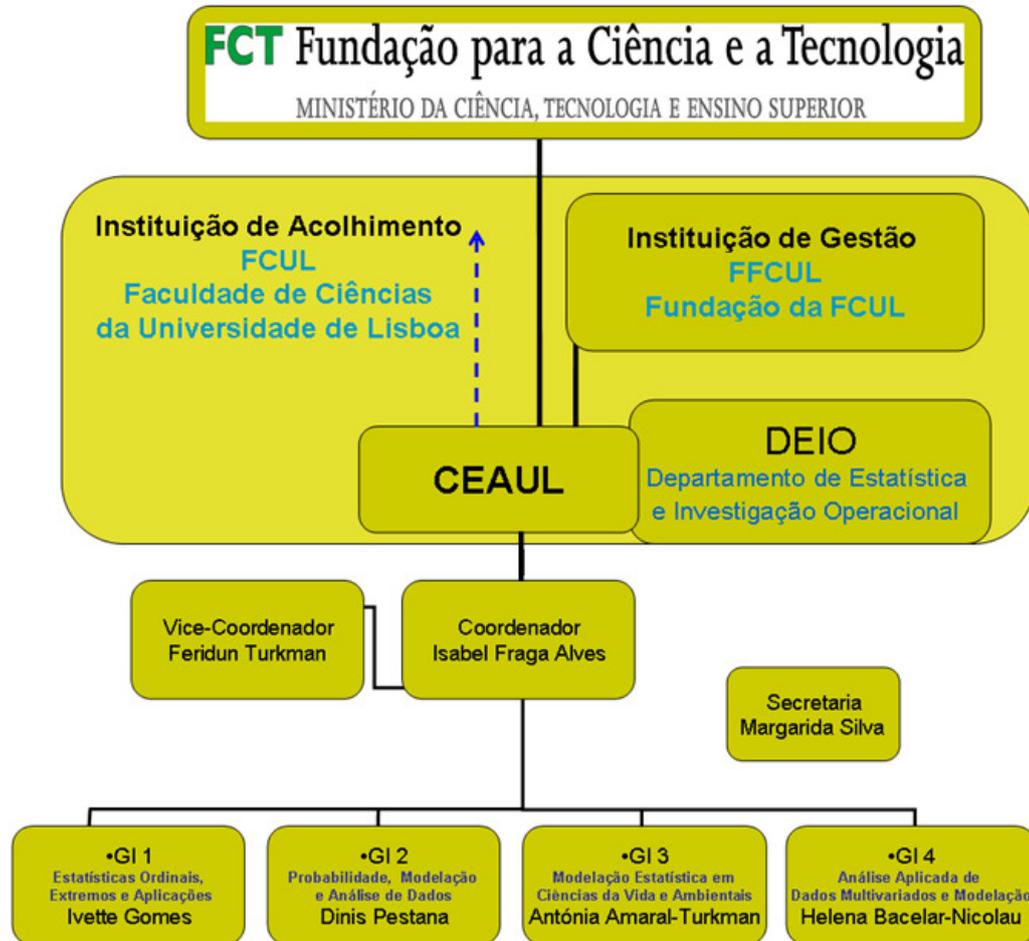
O CEAUL é membro associado do CIM-Centro Internacional de Matemática/ ERCOM - European Research Centres on Mathematics – Portugal.

A organização interna do CEAUL está sistematizada em 4 grupos de investigação (GI's), partilhando interesses científicos similares em Probabilidade, Inferência Estatística, Modelação e Análise de Dados, mas com ênfase diverso relativamente a metodologias de aplicação segundo plano director de cada Investigador Principal. Os GI's são encorajados a interagir nas suas actividades, sempre que tal seja pertinente. Os GI's com que o CEAUL concorreu para o financiamento plurianual para 2007-2010 são:

- GI 1: Estatísticas Ordinais, Extremos e Aplicações;
- GI 2: Probabilidade, Modelação e Análise de Dados;
- GI 3: Modelação Estatística em Ciências da Vida e Ambientais;
- GI 4: Análise Aplicada de Dados Multivariados e Modelação.

2. Descrição sumária do CEAUL

O enquadramento desta unidade de I&D situa-se num binário FCT / FCUL, se bem que com esta última acresce uma situação de mudança estrutural profunda, fruto da implementação para breve dos novos Estatutos da Faculdade.



O CEAUL visa contribuir para o desenvolvimento e inovação do *background* teórico e das metodologias dirigidas às aplicações aos problemas reais, fazendo esforços de forma a corrigir o uso da Estatística, quer na Sociedade quer no meio Académico. Este objectivo tem vindo a ser alcançado com sucesso através de diferentes actividades interligadas: conferências e seminários, publicações científicas em revistas de peritagem nacional e internacional da especialidade, Pós-Graduação, a actualização da Biblioteca, que representa um caso ímpar em Portugal na área da Estatística, aberta não só ao meio Académico como à Comunidade, em geral.

As áreas em que o CEAUL investiga são de largo espectro e concorrem com os seus pares a nível internacional em:

- Health Sciences and Biology
- Natural Hazards in Geophysics and Environment
- Computational Statistics and Simulation
- Quality Control
- Biostatistics
- Survival Analysis
- Time Series
- Sampling
- Automatic Classification
- Design of Experiments
- Financial Risk
- Statistics of Extremes
- Resampling Methodologies
- Bayesian Statistics
- Genomics and Proteomics
- Clinical Trials
- Data Analysis
- Psychology
- Temporal and Spatial Statistics
- Teaching of Statistics

O CEAUL está sistematizado em 4 Grupos de Investigação (GI), segundo plano de cada Investigador Principal. Os líderes de 4 grupos, junto com o Coordenador eleito e o Vice-Coordenador, são o núcleo do Conselho Científico cuja responsabilidade é não só monitorizar os objectivos principais da unidade como estabelecer a estratégia política de desenvolvimento para governar as políticas de investimento. O domínio principal do GI1 reside na área dos Valores Extremos, Controlo de Qualidade, parâmetros de Acontecimentos Raros ou Catastróficos e Risco; GI2 foca a sua acção em Modelos Probabilísticos, com Aplicação às Ciências Sociais e da Vida, e Ensino da Estatística. Os objectivos de GI3 estão direccionados para o tratamento de Dados Ambientais e Biológicos e Modelação. GI4 está mais vocacionado na Análise da Modelação Multivariada, *data mining*, em Ciências Comportamentais, Educação e Biomedicina. No quadro abaixo são referidas as áreas que prioritariamente são exploradas por cada Grupo, com as suas designações em Inglês.

GI 1 Estatísticas Ordinais, Extremos e Aplicações



Extreme Values
Tail behaviour
Modelling Catastrophic or Rare Events
Statistical Process Control Quality

Financial Risk
Risk Assessment
Environmental Risk



GI 2 Probabilidade, Modelação e Análise de Dados



Stopped sums and extremes
Probability models
Teaching Statistics
Experimental Design
Longitudinal data
Small sample asymptotics
Modelling biological and linguistic data

GI 3 Modelação Estatística em Ciências da Vida e Ambientais



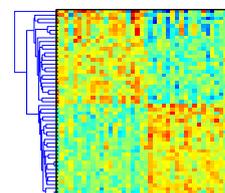
Modelling times of occurrence and size of forest fires
Prediction of local precipitation in southern Portugal through empirical spatio-temporal models
Modelling environmental and biological data

Genomics and Proteomics
Spatial and temporal Processes
Bayesian Methodology
Survival Analysis
Statistics in Genetics



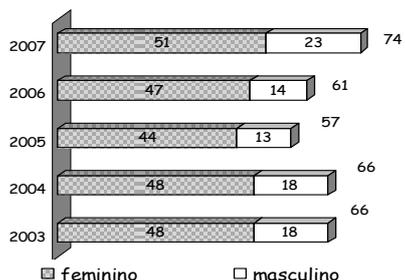
GI 4 Análise Aplicada de Dados Multivariados e Modelação

Classification and Data Mining in Behavioural Sciences
Cluster analysis
Education and Biomedicine
Missing Data
Experimental Design Sampling

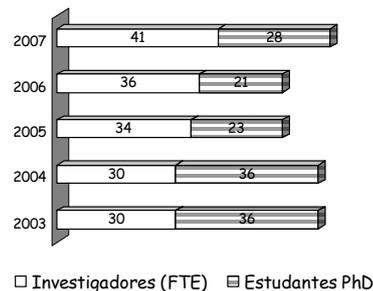


3. CEAUL em Números

(#) CEAUL – equipa ♂ vs. ♀



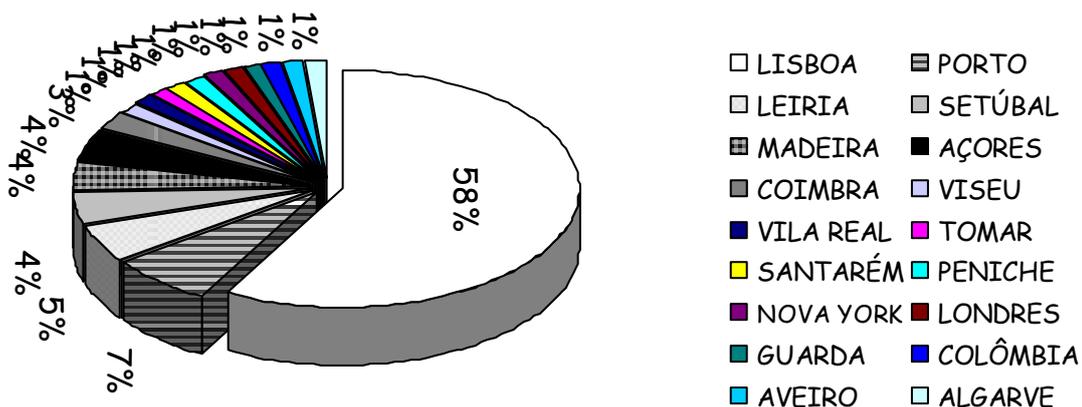
(#) Investigadores FTE vs. Estudantes PhD



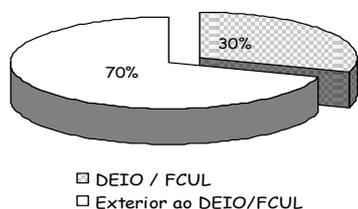
Os estudantes-PhD aqui contabilizados são “colaboradores” do CEAUL; os estudantes-PhD orientados por membros do CEAUL são maioritariamente de outras instituições fora do DEIO; frequentemente estes estudantes estão não só a exercer actividade docente com elevado número de horas lectivas, mas também acumulam funções burocráticas de enorme peso e asfixia para a sua investigação; nesta situação os Institutos Politécnicos constituem um caso “gritante”, que esperamos em parte seja superado com o recente Programa FCT/PROTEC – (Programa de apoio à formação avançada de docentes do Ensino Superior Politécnico).

(%) CEAUL por Regiões

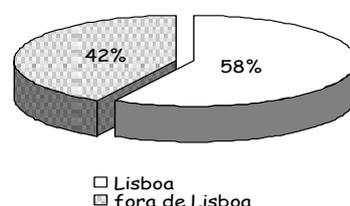
Actualmente a equipa de CEAUL, actualizada em 2008, conta com 43 doutorados integrados com 46 membros colaboradores constituindo uma gama muito diversificada, quer de instituição de proveniência quer no leque de competências mais específicas, dando cobertura a nível nacional. Como corolário, a unidade desempenha papel central em Portugal como centro de investigação fundamental e aplicada nas áreas da Probabilidade e da Estatística.



(%) DEIO/FCUL vs. fora de DEIO/FCUL

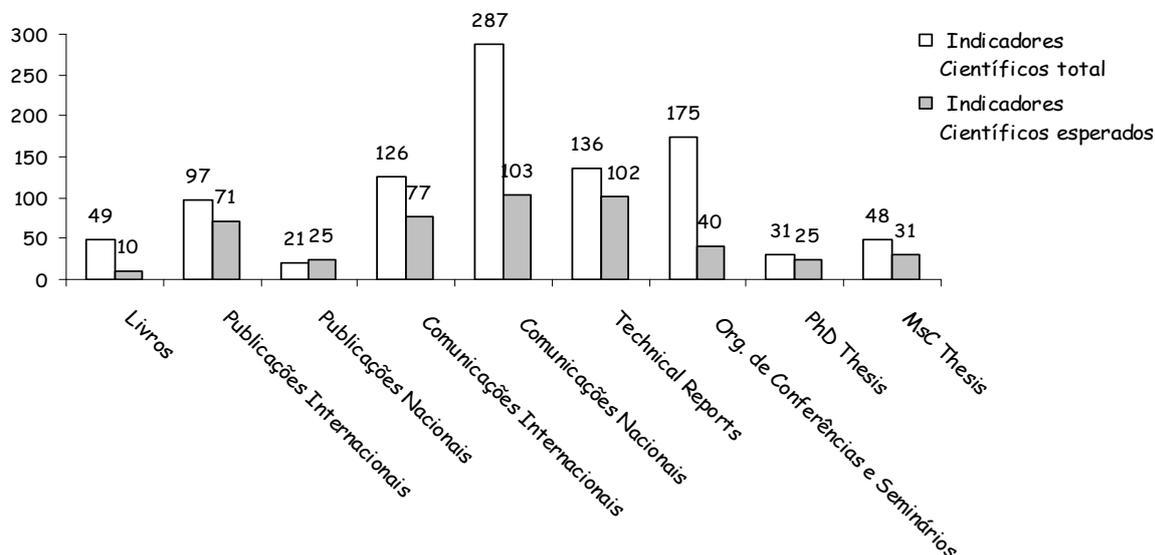


(%) Área Lisboa vs. fora de Área Lisboa



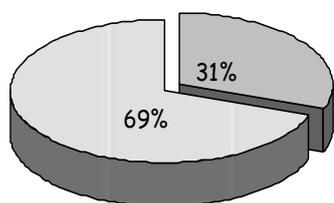
A política estratégica do CEAUL como um todo, para além da intensa publicação científica em revistas de qualidade e proeminente representação em Comunicações em Conferências a nível internacional e nacional, abrange ainda um forte investimento na comunicação do interesse da Estatística à Sociedade em geral e atraindo estudantes para uma área onde a taxa de empregabilidade é elevada.

(#) CEAUL - Indicadores Científicos (Esperados vs Observados 2003-2006)



Comunicações Internacionais vs. Nacionais

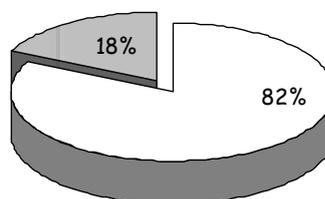
(Conferências, Workshops, etc.)



■ Comunicações Internacionais
□ Comunicações Nacionais

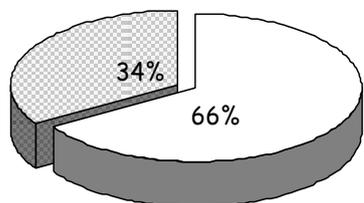
Publicações Internacionais vs. Nacionais

(Revistas)



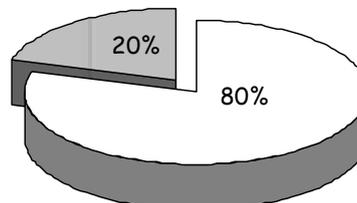
□ Publicações Internacionais
■ Publicações Nacionais

Publicações Internacionais



□ Revistas Internacionais JCR2006 ISI Web of Knowledge
□ Outras peer review Revistas Internacionais

Publicações Internacionais

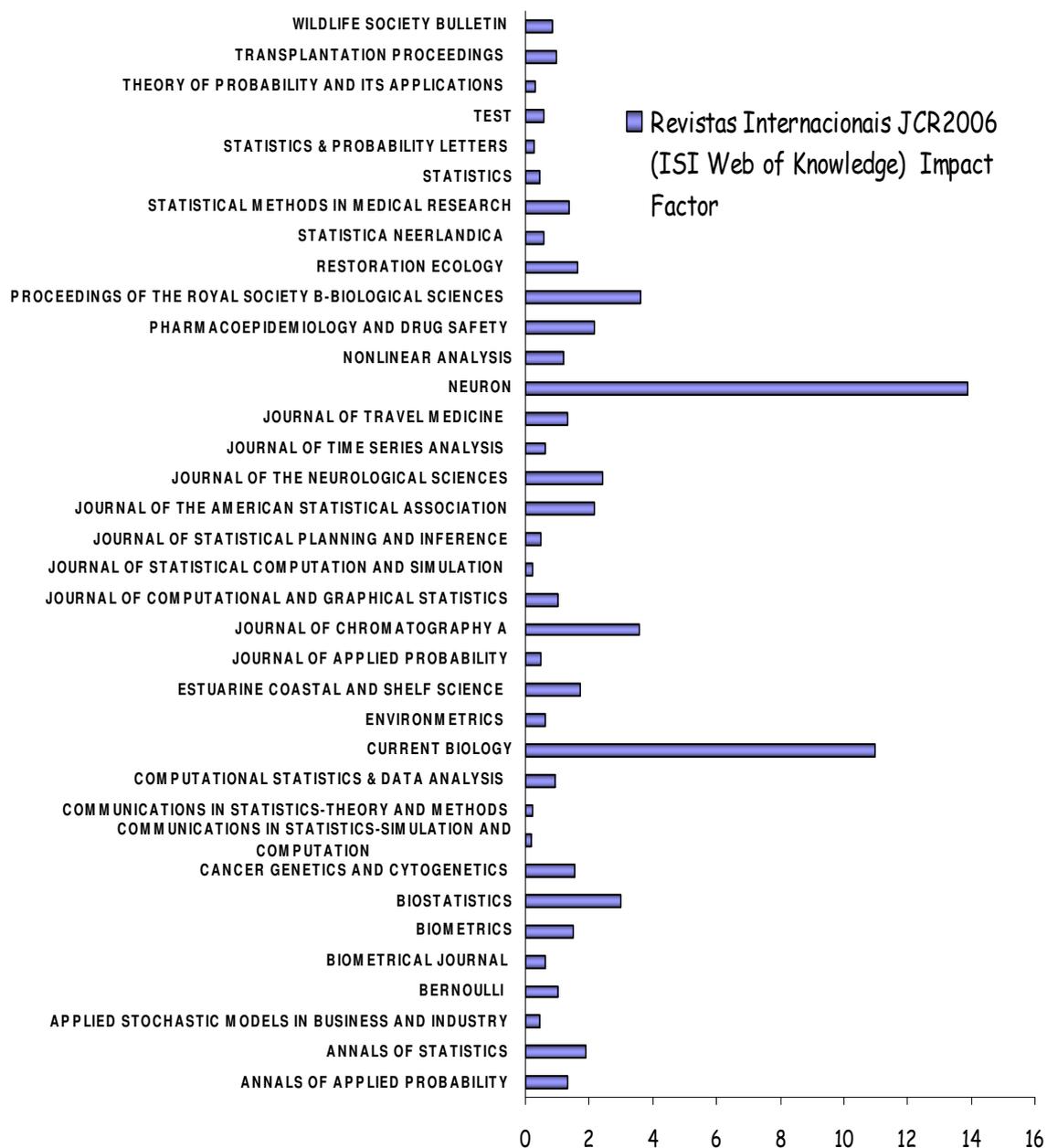


□ JCR2006 ISI Web Maths &al.
■ JCR2006 ISI Web Bios &al.

Os membros do CEAUL vêm os seus trabalhos publicados em revistas científicas de especialidade: Annals of Applied Probability; Annales de L'IS.U.P; Annals of Statistics; App. Stoch. Models in Business and Industry; Bernoulli; Biometrical Journal; Biometrical Letters; Biometrics; Biostatistics; Cancer Genetics and Cytogenetics; Coastal and Shelf Science; Colloquium Biometryczne; Communications in Statistics — Simulation and Computation; Communications in Statistics — Theory and Methods; Computational Statistics and Data Analysis; Current Biology; Environmetrics; Extremes; Human Resources for Health; Int. J. Comm. in Depend. Quality Management; International J. of

Ichthyology; J. American Statistical Association; J. Appl. Probab.; J. Computational and Graphical Statistics; J. Soc. Esp. del Rorschach y Mét. Proyectivos; J. Statist. Comp. and Simulation; J. Statist. Planning and Inference; Journal of Chromatography A; Journal of Infection, Genetics and Evolution; Journal of Neurological Sciences; Journal of Nonlinear Analysis; Journal of Time Series Analysis; Journal of Travel Medicine; Mathematical Methods of Statistics; Neuron; Parassitologia; Pharmacoepidemiology and Drug Safety; Portugaliae Mathematica; Proc. Royal Soc. London B; Publications de l'Institut Mathématique; Restoration Ecology; RevStat; Statistica Neerlandica; Statistical Methods in Medical Research; Statistics; Statistics and Probab. Letters.; Test; Theory Probab. and its Applications; Transplantation Proceedings; Wildlife Society Bulletin.

Embora o Impact Factor (ISI Web Knowledge) constitua uma unidade questionável, referenciamos aqui a ordem de grandeza daquele factor, que claramente apresenta escalas diferentes consoante se trate de áreas ditas mais “matematizadas” com as de índole mais aplicada às ciências da vida e afins, e penalizando fortemente as primeiras.



A interação do CEAUL com grupos paralelos a nível internacional é uma nota presente desde sempre, sendo evidenciada pela forte publicação em colaboração com cientistas estrangeiros que frequentemente visitam o CEAUL e de forma simétrica convidam membros do CEAUL para uma interação profícua e com resultados relevantes. O CEAUL tem papel dinâmico e interventivo na realização de seminários regulares e de encontros mais alargados a nível nacional e internacional. Nos últimos 5 anos, foi possível contar com alguns cientistas de renome mundial: Alan Gelfand (Duke

University, USA); Andrei Khrennikov (Vaxjo, Sweden); Anne-Laure Fougères (Toulouse, France); Anthony Davison (Lausanne, Switzerland) ; Armelle Guillou (Paris, France); Byung Soo Kim (Seoul, S. Korea); Chris Cannings (Sheffield, UK); Cliff Spiegelman (College Station, USA); Clive Anderson (Sheffield, UK); Dave Allcroft (Scotland, UK); David Banks (Durham, USA); David Ríos Insua (Madrid, Spain); Dirk Husmeier (Scotland, UK); Doug Nychka (USA); Elisaveta Pancheva (Bulgary); Filip Lindskog (Stockholm, Sweden); Fred Smith (UK); Gyorf Laszlo (Budapest, Hungary); Hélio S. Migon (Rio de Janeiro, Brasil); Hildete Prisco Pinheiro (São Paulo, BRASIL); Holger Drees (Saarland, Germany); Holger Rootzen (Gothenburg, Sweden); Jan Beirlant (Leuven, Belgium); Jef Teugels (Leuven, Belgium); Jef Teugels (Leuven, Belgium); Jim Zidek (British Columbia, Canada); Johan Segers (Tilburg, The Netherlands); John Einmahl (Tilburg, The Netherlands); Jonathan Tawn (Lancaster, United Kingdom); Korbinian Strimmer (Munich, Germany); Laurens de Haan (Lisbon, Portugal & Rotterdam, The Netherlands) ; Leonhard Held (Munich, Germany); Liang Peng (Atlanta, USA); Manuela Zucknick (London, UK); Maria Ivanilde Araújo (Amazonas, Brasil); Martin Schlather (Göttingen, Germany); Michael Falk (Wuerzburg, Germany); Nathalie Thorne (Cambridge, UK); Noel Cressie (Ohio, USA); Nozer Singpurwalla (Washington, D.C., USA); Patrick Brown (Lancaster, UK); Paul Blackwell (Sheffield, UK) ; Paulo Justiano Ribeiro (Paraná, Brazil); Peter Hall (Australian National University); Philippe Broët (Paris, France); Philippe Naveau (Gif-sur-Yvette, France) ; Piet Groeneboom (Delft and Amsterdam, The Netherlands); Pieter van Gelder (Delft, The Netherlands) ; Richard Davis (Colorado, USA); Richard Kay (UK); Rolf Reiss (Siegen, Germany); Ross Leadbetter (North Carolina, Chapel Hill, USA); Ruedi Aebersold (Seattle, USA); Samorodnitsky, Gennady, Cornell University; Sidney Resnick (Cornell, USA); Simon Tavaré (South California, USA); Soeren, Asmussen (Aarhus, Denmark); Sophie Schbath (Jouy-en-Josas, France); Sreenivasan Ravi (Mysore, India); Stilian Stoev (Michigan, USA); Sujit Sahu (Southampton, UK); Terry Speed (California, USA); Thomas Mikosch (University of Copenhagen); Tilmann Gneiding (Washington, USA) ; Valerie Isham (London, UK); Vladas Pipiras (North Carolina, Chapel Hill, USA); Vladimir Piterbarg (Moscow, Russia); Wolfgang Urfer (Dortmund, Germany).

A nível nacional, refira-se o lugar de destaque de colaboração estreita do CEAUL com a SPE, ocupando lugar de destaque o elevado número de participantes na Conferência anual daquela Sociedade, acrescido de um relevante número de publicações nas Actas daquelas Conferências. A génese da SPE(IO) confunde-se com o CEAUL, uma vez que na sua origem estiveram elementos fundadores deste Centro. Actualmente, CEAUL e SPE partilham recursos humanos e materiais na FCUL.

O CEAUL tem ainda cooperado com o INE, através do envolvimento de alguns dos seus membros séniores; em 2003, foi concretizado o projecto da implementação e produção da REVSTAT – Statistical Journal, uma Revista Internacional na área da Probabilidade e da Estatística [1 vol. (= 3 números) por ano], <http://www.ine.pt/revstat/inicio.html>, editada pelo INE, constituindo já um êxito no seio da academia científica nacional e internacional; outro projecto com o INE foi a organização em Lisboa da maior Conferência Internacional na área, *56th Session International Statistical Institute* (2007), sendo a Presidente do Programa Local a Prof. Ivette Gomes, grande impulsionadora do CEAUL e prévia Coordenadora da unidade.

4. Avaliação do CEAUL

A avaliação referente ao período 2003-2006 foi de Muito Bom. O resultado foi fruto de um relatório elaborado por um Painel de Avaliação Internacional que visitou o nosso Centro em Fevereiro de 2008. Note-se que esta unidade desempenha papel central em Portugal como centro de investigação fundamental e aplicada nas áreas da Probabilidade e da Estatística, não existindo em Portugal outra unidade de I&D comparável a nível destas características. O processo de avaliação na área da Matemática revestiu-se de facetas bizarras: por um lado, do painel apenas consta um especialista na área específica da Estatística; por outro lado, houve duas épocas distintas de avaliação (Fevereiro e Julho de 2008), diferindo a constituição do Painel; no caso particular da área da Estatística, o perito internacional da 1ª ronda foi o Prof. Richard Smith, substituído na 2ª ronda pelo Prof. Anthony Davison. Qualquer avaliação comparativa entre as unidades de I&D portuguesas na área da Matemática/Estatística falha já que assenta em pressupostos distintos.

Este tipo de situações serão obviamente de evitar como facilmente qualquer planeamento de experiências advverte.

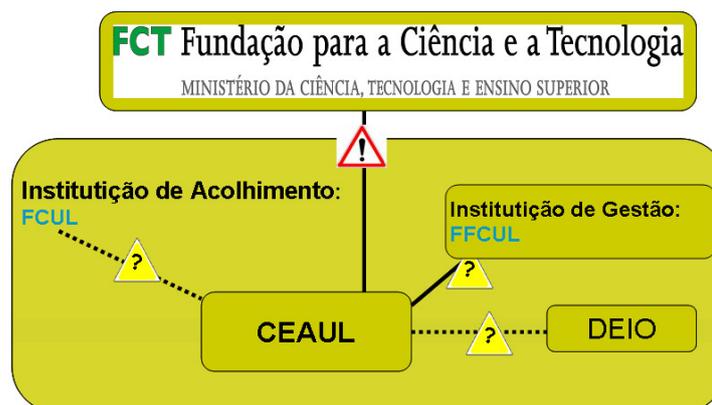
5. Futura visão da estratégia do CEAUL

Além das metas que o CEAUL tem vindo a atingir, refiram-se ainda:

- **CEAUL(FCUL) / Unidades I&D da UL:**
Formação de Integração na Investigação através do Programa BII/Ciência2008 em colaboração com outros Centros da Universidade de Lisboa;
- **CEAUL(FCUL) / Empresa:**
Formação Avançada (PhD) em ambiente empresarial, sob supervisão de ambos os lados;
- **CEAUL(UL) / Instituto Gulbenkian de Ciência:**
CEAUL é interveniente no Consortium de Epidemiologia de Lisboa – UL / IGC;
- **CEAUL(FCUL) / INE:**
projecto conjunto com o INE - Novas Tabelas de Mortalidade Portuguesas;
- **CEAUL / ALEA-INE:**
Programa de partilha de experiências de ensino com professores das escolas secundárias;
- **CEAUL / Sociedade:**
Promoção da Estatística na sociedade, cidadania e comunicação social.

6. Maiores Dificuldades do CEAUL

O esquema seguinte sumariza as principais apreensões do CEAUL à presente data.



⚠ FCT / CEAUL-FCUL

Os Termos de Aceitação FCT/CEAUL(FFCUL) para o programa plurianual, têm vindo sistematicamente a constituir contratos escritos não fiáveis. No que me diz respeito como Coordenadora do CEAUL, acrescento ainda a dificuldade acrescida de calendarizar tarefas estrategicamente programadas, quando existe um desfazamento total entre os prazos fixados pela entidade financiadora e o incumprimento real dos pressupostos nos contratos iniciais. Coordenar este Centro de Investigação tem-se vindo a revestir de características de uma “Missão Impossível”. Não diria que existe falta de apoio ao investigador, mas sim uma gestão não otimizada dos meios para um saudável desenvolvimento científico. Urge agilizar a pesada infra-estrutura de apoio à Investigação e alterar uma visão global de todo um processo que se reveste de um princípio básico de que as unidades de I&D e seus investigadores não constituem “Pessoas de Bem”. O Investigador tem de ser acreditado com seriedade, aumentando a sua auto-estima e contribuindo para um avanço real da investigação em Portugal. Infelizmente não é esta a situação, quer por parte da Instituição Financiadora quer por parte da Instituição de Gestão. Como classificar o facto de a FCT proceder ao pagamento do 2º semestre de 2007 nos últimos dias de Dezembro desse ano, contrariando totalmente os pressupostos que constam do respectivo Termo de Aceitação com o CEAUL? Qual o sentido de equidade que preside a essa decisão tomada apenas para alguns Centros, em total desarmonia com outras unidades de I&D a quem foi cumprido o prazo de pagamento em Julho de 2007? Que justiça está subjacente quando a FFCUL pretende aumentar os overheads de 4% para 20% em 2008 sobre os montantes referidos, penalizando o

CEAUL mais uma vez pelo facto da entidade financiadora proceder de modo desigual relativamente aos plurianuais das outras unidades de I&D ? Como se explica o facto da FCT ter em 2009 procedido a pagamentos avultados relativos a verbas em falta de 2006, numa altura de crise económica?

Não querendo repetir-me, realmente considero lamentável a forma como é gerido o dinheiro público por parte da FCT, como se está a revelar, entre outros, o caso das Bolsas BII. Efectivamente, a 12 de Maio de 2008, foi lançada com *pompa e circunstância* na Universidade de Aveiro, a iniciativa Ciência 2008, da qual constava o novo Programa de Bolsas de Integração na Investigação (BII). Na sequência do “Compromisso”(?) com a Ciência “assumido” pelo Governo e nos termos do Regulamento de Formação Avançada e Qualificação de Recursos Humanos, a FCT convidou as Instituições de Investigação do Sistema Científico e Tecnológico Nacional (SCTN) a abrirem (entre Maio e Outubro) concurso para atribuição BII a serem financiadas pela FCT, com o objectivo de virem a ser atribuídas até 5000 bolsas de um ano (140 Eur/mês) em 2008, perfazendo um total 8.400.000 Eur. O CEAUL viu neste programa a oportunidade de concretizar uma meta estratégica de, não só fazer despertar nos alunos de 1º ciclo o interesse pela investigação, mas aliar a este programa uma cooperação entre Centros de Investigação da UL. Lançado este repto BII/CEAUL/2008, foi com prazer que vimos igualmente abraçada esta ideia com interesse e a postura séria inerente ao desenhar de projectos comuns para o ano lectivo 2008/09. Tudo se passou com um grande investimento na esperança que desta vez esse “Compromisso”(?) com a Ciência fosse realmente “assumido” pelo Governo! *Hélas!* Decorridos os prazos estipulados para as unidades procederem de acordo com *1001* burocracias, eis que estando a iniciar o 2º semestre e ainda não se vislumbra qualquer tipo de resposta por parte da FCT no sentido de aprovação ou não da bolsa a que concorremos.

Será este o procedimento adequado para despertar nos jovens o interesse pela investigação?

Ou será ao invés, uma forma de desde cedo a juventude do nosso país constatar como se consegue defraudar as expectativas e mais uma vez ser gorada a esperança de que as Instituições governamentais “assumam efectivamente os seus compromissos”.

Quando vier uma eventual resposta por parte da FCT, o aluno pode já não ter possibilidades de aceitar a bolsa, pois o ano lectivo 2008/09 já passou! E para os Centros reparo análogo tem cabimento. *SEM PALAVRAS!*

Acrescem ainda ao rol das dificuldades de gestão do CEAUL outros factos inqualificáveis. As despesas inerentes às actividades das unidades de I&D são muitas das vezes catalogadas de “não elegíveis” aos “olhos” da FCT. *A priori* parece que tudo se passa como se os Investigadores tenham obrigação de provar que são “gente honesta” com boas intenções, o que produz um desgaste permanente e uma canalização de energias que deveriam ser preservadas para uma investigação mais profícua. Gostaria ainda de acrescentar que urge ser implementada a medida, preconizada no “*Compromisso com a Ciência*” do MCTES, de “*Gestão diferenciada do tempo lectivo de cada docente, em função da sua actividade de investigação*”: a realidade portuguesa enquadra a Investigação paralelamente à Docência como obrigações concomitantes na Carreira Docente do Ensino Superior.



Por outro lado, a “relação” UL ↔ FCUL ↔ FFCUL ↔ CEAUL ↔ DEIO tem sido ambígua até agora: embora a maioria das unidades I&D estejam alojadas em Universidades/Faculdades, caso do CEAUL, não existia até agora um enquadramento nos regulamentos da Instituição de Ensino, afigurando-se os Centros mais a entidades “*fantasmas*” que frutificavam num mundo paralelo. É com satisfação que constato que os novos **Estatutos da Faculdade de Ciências da Universidade de Lisboa** (DR: 6 Fev 2009), elaborados de acordo com os Estatutos da Universidade à luz do RJIES, contemplam finalmente a figura de *Centros de Investigação associados à FCUL*, com direitos e deveres inerentes. É este o caso do CEAUL de acordo com ANEXO C, alínea 6).

Estatutos da Faculdade de Ciências da Universidade de Lisboa

Artigo 10.º

Natureza das Unidades de Investigação e Desenvolvimento

1 — As *unidades de investigação e desenvolvimento* são entidades reconhecidas e financiadas pelo sistema científico e tecnológico nacional, dotadas de autonomia científica, que integram docentes ou investigadores da Faculdade e ou

docentes ou investigadores externos, as quais desenvolvem actividade de investigação sob gestão da Universidade, da Faculdade, ou de instituições por elas participadas.

2 — As *unidades de investigação e desenvolvimento* estão associadas a pelo menos um dos departamentos com que partilham recursos humanos e materiais.

3 — As *unidades de investigação e desenvolvimento* são necessariamente coordenadas por docentes ou investigadores com vínculo contratual com a Universidade ou com a Faculdade.

Conselho Científico

Artigo 31.º

Função

O conselho científico é o órgão de natureza científica e cultural bem como de planeamento estratégico da Faculdade.

Artigo 32.º

Composição

1 — O conselho científico é composto pelo Director que preside e 24 professores e investigadores, assim distribuídos:

- a) 14 professores ou investigadores doutorados com vínculo à Faculdade ou à Universidade, sendo que, pelo menos, quatro quintos devem ser professores ou investigadores de carreira;
- b) *10 representantes das unidades de investigação* com vínculo à Faculdade ou à Universidade.

Estando a completar o meu mandato de Coordenadora do CEAUL, permitam-me uma última palavra de alento para a coordenação vindoura: perante este excerto dos novos Estatutos da FCUL, gostaria de acreditar que, num futuro que se adivinha próximo, a cooperação efectiva com a FCUL, e com o DEIO em particular, permitirá concretizar metas óbvias decorrentes das valências do Centro, envolvendo o fortalecimento da cooperação científica com outros Departamentos e unidades de I&D da FCUL. É sem dúvida desejável que se efective uma abertura a parcerias que potenciem a oferta de programas inovadores que contribuam para reforçar a capacidade científica e de formação avançada na FCUL.



Centro de Estatística e Aplicações
Universidade de Lisboa

<http://ceaul.fc.ul.pt>



Algumas Reflexões Avulsas sobre a Sociedade Portuguesa de Estatística

M. Ivette Gomes, *ivette.gomes@fc.ul.pt*

Universidade de Lisboa, FCUL/DEIO e CEAUL

Apesar da minha fraca memória para relembrar factos históricos, vou procurar referir brevemente alguns dos pioneiros no desenvolvimento da *Estatística* em Portugal nestes últimos 40 anos e proceder a algumas reflexões avulsas sobre a *Sociedade Portuguesa de Estatística*.

Nos dois últimos anos da minha licenciatura em *Matemática Pura*, especialidade de *Álgebra*, senti-me fascinada pelo pouco que tinha aprendido sobre *Estatística*, termo considerado aqui em sentido lato e englobando a *Probabilidade* e os *Processos Estocásticos*. Apesar de um convite aliciante do Professor Almeida Costa, eminente algebrista da *Faculdade de Ciências da Universidade de Lisboa (FCUL)*, para ir directamente para os Estados Unidos da América (EUA), com uma bolsa de estudos da Gulbenkian, tomei a decisão de ir dar aulas para o colégio de um familiar e de continuar a estudar, de modo a poder obter uma licenciatura em *Matemática Aplicada* e se possível enveredar mais tarde por investigação em temas da área da *Estatística*. Esta minha tomada de decisão chegou, por portas e travessas, aos ouvidos do Professor Tiago de Oliveira, professor do *Departamento de Matemática* da *FCUL* e estatístico de renome internacional, que de imediato me contactou, convencendo-me a entrar como Assistente da *Secção de Matemática Aplicada* da *FCUL*, onde poderia estudar os temas que me agradavam numa categoria de “aluno professor”. Nessa altura existiam poucos nomes sonantes nesta área. Em Lisboa, recordo os Professores Pedro *Braumann*, M. Fátima *Fontes de Sousa*, Bento *Murteira* e José *Tiago de Oliveira*. Havia também o Professor Manuel *Murta*, em Coimbra, de que me lembro mal, e o Professor Diogo *Pacheco de Amorim*, que não conheci. Mais novo, já doutorado e trabalhando na área, lembro-me apenas do Professor *Galvão de Melo*. Não posso ainda deixar de referir académicos da área da *Análise*, que também fizeram trabalho meritório em *Estatística*, tais como os Professores *Dias Agudo*, *Passos Coelho* e António *St Aubyn*, tendo este último dedicado desde há muito a sua investigação à área da *Estatística*.

Houve então, a partir do início dos anos 70, uma pressão generalizada para a saída do País dos jovens cientistas da altura, de modo a obterem os seus graus no estrangeiro, essencialmente em França, em Inglaterra ou nos EUA. Na década de 70 partiram para França a Helena *Barroso*, a Lucília *Carvalho*, a Eugénia *Graça Martins*, a Nazaré *Mendes-Lopes*, o Daniel *Muller*, o Fernando *Nicolau* e a Helena *Nicolau*, certamente por entre outros. Para Inglaterra foram a Antónia *Amaral*, o João *Branco*, o Dinis *Pestana*, a Fernanda *Ramalhoto*, a Cristina *Sernadas*, e eu própria, certamente também por entre outros. Nessa mesma época e para os EUA só me recordo de ter ido o Carlos *Braumann*. Todos os que foram para Inglaterra ou EUA voltaram para Portugal com o grau de Doutor no fim da década de 70 ou inícios da década de 80, e tiveram equivalência ao grau de Doutor em Portugal. Não tiveram sorte idêntica os que foram para França, que voltaram para Portugal com o “*Troisième Cycle*”, a que não era dada equivalência a Doutoramento, e que tiveram de efectuar provas de Doutoramento em Portugal. Mas à excepção da minha colega e amiga Helena *Barroso*, que não continuou carreira universitária, todos os outros se doutoraram até finais da década de 80.

Saíram mais tarde, já na década de 80 e com a obtenção do grau de Doutor até fins de 1990, o João *Amaral*, a Margarida *Brito*, a Lourdes *Centeno*, o Paulo *Gomes*, a Esmeralda *Gonçalves*, a Manuela

Magalhães, o Álvaro Nunes, o Carlos Daniel Paulino, a Elizabeth Reis, a M. Carolina Silva e a Corália Vicente, também por entre outros. E continuámos sempre a ter doutoramentos no estrangeiro, mas com um aumento significativo dos doutoramentos no País. E recordo aqui a frase da Antónia Amaral-Turkman no livro editado pelo Fernando Rosado (2005), *Memorial da Sociedade Portuguesa de Estatística*: “dos 139 doutoramentos contabilizados desde 1977 até Fevereiro de 2005, 96 foram realizados em Portugal e 43 no Estrangeiro”. Mas, com base nos dados desse mesmo artigo, penso que dos 37 doutoramentos anteriores a 1990 (inclusive) só devem ser contabilizados como doutoramentos em Portugal os de M. Teresa Alpuim, Salomé Cabral, Francisco Calheiros, Helena Iglésias Pereira, Fernando Macedo, Beatriz Matias, João Tiago Mexia, M. Manuela Neves, José Rodrigues Dias e Fernando Rosado. Acrescento pois que até fins de 1990, a percentagem de doutoramentos realizados (total ou parcialmente) no Estrangeiro foi superior a 70%, sendo essa mesma percentagem inferior a 25% para doutoramentos obtidos a partir de 1991, o que sem dúvida confirma a instalação de “know-how” suficiente no País, na área da Estatística.

Eu e o Dinis Pestana regressámos a Lisboa, já doutorados, em fins de 1978, ainda para a *Secção de Matemática Aplicada* do *Departamento de Matemática* da *FCUL*, e ambos assistimos a dois dos marcos importantes para o desenvolvimento da *Estatística* em Portugal. Um desses marcos foi sem dúvida a fundação da *Sociedade Portuguesa de Estatística e Investigação Operacional (SPEIO)*, em 1980 (Novembro, 28), com onze outorgantes da escritura de formação (J. Tiago de Oliveira, J. Dias Coelho, Pedro Braumann, Bento Murteira, M. Fernanda Ramalhoto, M. Ivette Gomes, Dinis Pestana, M. Cristina Sernadas, Amílcar Sernadas, M. Antónia Amaral-Turkman e M. Fátima Fontes de Sousa) e poucas dezenas de sócios fundadores (todos os que aderiram até 31 de Julho de 1981). O outro marco foi a criação do *Departamento de Estatística, Investigação Operacional e Computação (DEIOC)*, em 1981. Tive a sorte de ser elemento activo quer na constituição do *DEIOC* quer na constituição da *SPEIO*, embora não possa deixar de referir o papel primordial e pioneiro de J. Tiago de Oliveira.

No âmbito da *SPEIO*, realizaram-se então dois *Congressos de Estatística e Investigação Operacional*, um no Fundão/Covilhã, em 1981, e outro, em Lagos, em 1984. Mas face à existência da *Associação Portuguesa para o Desenvolvimento da Investigação Operacional (APDIO)*, em que colaboravam activamente os sócios da *SPEIO* da área da *Investigação Operacional*, a sensibilidade para uma mudança de nome da *SPEIO* para *SPE (Sociedade Portuguesa de Estatística)* foi-se instalando na comunidade.

A partir de 1989 procedeu-se a uma reestruturação da *SPEIO*, em que estive profundamente envolvida (juntamente com M. Antónia Amaral-Turkman e António Simões Neto, entre outros). Num período conturbado, realizou-se uma 1ª *Conferência em Estatística e Optimização*, em Tróia, 1990, promovida não pela *SPEIO/SPE*, mas pelo *Centro de Estatística e Aplicações da Universidade de Lisboa (CEAUL)* e pelo *DEIOC/FCUL*. Em 1991 foram aprovados novos estatutos e adoptado o presente nome de *Sociedade Portuguesa de Estatística (SPE)*. Estive à frente da *SPEIO/SPE* desde 1989, assumindo a Presidência da *SPE* durante o período 1990-1994. É pois com grande satisfação que tenho visto a *SPE* crescer e afirmar-se. Ao tempo em que presidi à *SPE* punha-se a tarefa urgente de reformular os seus estatutos, uma tarefa aparentemente simples mas que limitou a capacidade de realizar outras tarefas, porventura mais nobres. Registo no entanto, com satisfação, que se deu início aos *Congressos Anuais* da *SPE*, que continuam a ser um ponto alto do calendário dos estatísticos em Portugal. E ainda me recordo das vozes dissonantes na *Assembleia Geral* da *SPE* em que se decidiu proceder à realização de um *Congresso Anual da Sociedade Portuguesa de Estatística*, com a marcação do primeiro Congresso, realizado no Hotel Golf Mar do Vimeiro, em 1993. Mas penso que ninguém duvida hoje da importância e necessidade de realização destes Congressos, que espero continuem com a mesma dinâmica no futuro.

Assisti, nas presidências subsequentes, a um crescimento substancial da *SPE*, e à implementação de muitas iniciativas importantes. Para além dos *Congressos Anuais*, a *SPE* patrocina agora a realização regular de conferências em colaboração com instituições do ensino superior. O reconhecimento internacional da *SPE* tem sido francamente incrementado, tendo sido a *SPE*, juntamente com o *Instituto Nacional de Estatística (INE)*, patrocinadora da 56th *Session of the International Statistical Institute*, que decorreu em Lisboa, em 2007. Existe na realidade com o *INE* uma parceria que só pode trazer vantagens às duas instituições. O *Boletim da Sociedade* atingiu um

estatuto exemplar, e além de veicular notícias de interesse para os sócios, publica artigos de opinião e institucionalizou já a interessante iniciativa de realizar “números temáticos”. Sobretudo, o rosto jovem da *SPE* é cada vez mais visível, na organização e participação nos congressos, e nos órgãos institucionais, o que prenuncia um futuro risonho.

Algumas outras iniciativas têm, a meu ver, que ser continuadas com persistência.

- Vejo grandes vantagens em reforçar a importância da *SPE* a nível peninsular, ibero-americano, e dos *PALOP*, porventura nesta última vertente criando um programa de bolsas de estudo para estudantes africanos que pretendam realizar pós-graduação em *Estatística* em Portugal (a Universidade da Madeira, neste aspecto, é porventura um exemplo a seguir).
- Há por outro lado que promover ainda mais a nova geração de estatísticos, incentivando-os por exemplo a organizarem *workshops* em Portugal, nomeadamente apoiando-os na procura das verbas necessárias para os internacionalizar.
- A parceria com o *INE* tem dado excelente frutos, e a *SPE* tem também relações institucionais fortes com a *Sociedade Portuguesa de Matemática* e a *Associação Portuguesa de Classificação e Análise de Dados (CLAD)*. Mas a *Estatística* é uma ferramenta cada vez mais usada por todos, e conviria porventura alargar os laços a outras sociedades científicas, como por exemplo a *Sociedade Portuguesa de Metrologia*.

A *Estatística* pode orgulhar-se de ser um *ponto de encontro de todas as ciências*. Não são poucos os exemplos de estatísticos cuja formação de base era outra, ou que se distinguiram como criadores de *Estatística* a par de obra científica de vulto noutra área. Mas isso não obsta a que seja hoje uma área do saber mal conhecida pela maior parte dos utilizadores, que muitas vezes têm da *Estatística* uma visão muito parcelar, e por vezes com lacunas que seriam graves num profissional desta área. Nesse aspecto, é minha opinião que a *SPE* tem um papel de responsabilidade a desempenhar, com diversas vertentes, nomeadamente:

- Opinar sobre o ensino da *Estatística*, a nível superior, em Portugal. Tal é imprescindível, naturalmente no que respeita à composição de cursos superiores em *Estatística*; mas também conviria ter alguma intervenção sobre o que se ensina, e como, a utilizadores de *Estatística*. Neste último aspecto, a acção junto de outras associações científicas pode ser um caminho privilegiado.
- Proceder sistematicamente à recensão crítica do que entre nós se vai publicando na área de *Estatística* e suas aplicações.
- Intervir na qualidade dos manuais escolares.

Há algumas razões para esperar - ou talvez para temer? - que uma Secção de *Ensino da Estatística* ganhe força, a curto prazo, dentro da *SPE*. Com essa área é necessário um cuidado redobrado, porque se realizações como o *ALEA* são um motivo de orgulho e esperança para todos nós, abundam no plano internacional (e porventura já entre nós) exemplos de essa área fundamental ser tomada de assalto por pessoas porventura bem intencionadas mas com uma fragilidade de conhecimentos preocupante.

Finalmente, face à estratégia de internacionalização aprovada na recente Assembleia da *SPE*, no que respeita a publicação em Inglês das *Actas dos Congressos da SPE*, é a meu ver uma tarefa de todos fazer crescer este *Boletim de Estatística*, nomeadamente reforçando os relatórios de trabalho em progresso mas já com resultados que mereçam divulgação. E por que não avançar com uma Secção semelhante ao “*Teacher's Corner*” do *American Statistician*, que serviria também para manter altos os padrões da área de investimento em *Ensino da Estatística*, não deixando que esta área fosse subalternizada a uma subárea das chamadas “*Ciências da Educação*”? E termino com o desafio de transformar o *Boletim da Sociedade Portuguesa de Estatística* numa revista nacional, com uma Secção de artigos sujeitos a recensão crítica dos pares.



Congressos da Sociedade Portuguesa de Estatística

Um barómetro da investigação em Estatística realizada em Portugal

João A. Branco, *joao.branco@math.ist.utl.pt*

*Centro de Matemática e Aplicações
Instituto Superior Técnico
Universidade Técnica de Lisboa*

1. Introdução

Os dezasseis congressos que a Sociedade Portuguesa de Estatística tem vindo a realizar desde 1993 até 2008 contêm informação preciosa que é indispensável ter em conta para responder a várias perguntas que se podem fazer sobre o desenvolvimento e a divulgação da Estatística em Portugal. Em particular faz todo o sentido perguntar o que se passa com a investigação científica em Estatística realizada no país, com vista a conhecer qual a natureza dessa investigação e a diversidade da mesma e qual o seu real valor, medido pela apreciação que a comunidade estatística faz desta actividade. Perguntas são fáceis de fazer mas, neste caso, obter respostas objectivas e claras é tarefa bem difícil. A informação com origem nos congressos, se bem que preciosa, é morosa de compilar e estudar e bastante complexa para ser assimilada com proveito. Por exemplo, no caso das actas o que está em jogo é um conjunto de 15 volumes com várias centenas de artigos e muitos milhares de páginas. Claro que o trabalho de responder às questões anteriores só pode ficar completo se forem consideradas as contribuições de outras fontes de informação que existam disponíveis. É o caso da publicação editada por Rosado (2005), Memorial da Sociedade Portuguesa de Estatística, um repositório de aspectos de grande relevo da história recente da Estatística em Portugal, que é afinal a história da estatística desenvolvida à volta da SPE. Outra fonte ainda são as actas e revistas da especialidade onde se publicam trabalhos directamente, sem passarem pelos congressos da SPE. Contudo não há a menor dúvida de que é dos congressos da SPE que brota o trabalho mais substancial produzido no país, sendo a garantia da sua qualidade assegurada pela supervisão cada vez mais exigente da SPE.

Neste contexto é justo declarar que ao falar deste tema não se pretende esgotar o assunto mas apenas abordar aqueles aspectos mais concretos, que o tempo permitiu trabalhar e que foram escolhidos para incluir nestas notas, aqui deixadas com humildade em face da complexidade e delicadeza da tarefa. Faz-se então o enquadramento do início da actividade de investigação em estatística e de seguida apresentam-se resultados de contagens sobre o conteúdo das actas, acompanhados de breves considerações, sonhando sempre que outros com interesse pelo tema as poderão continuar e com tempo as poderão estudar tirando ilações que certamente muitos de nós gostaríamos já de saborear.

2. Antecedentes

A Estatística considerada como actividade cujo objectivo é o de registar, organizar e descrever dados é muito antiga, enquanto que a Estatística encarada como disciplina científica que se ocupa do desenvolvimento de métodos e se preocupa com a sua aplicação na resolução de problemas do mundo real é relativamente recente. A Portugal chegaram a seu tempo estas duas visões da Estatística mas esse tempo é um tempo tardio se o compararmos com momentos idênticos que tiveram lugar naqueles países pioneiros das actividades estatísticas. Para tomar consciência da dimensão desse atraso basta, por exemplo, olhar para 1834, ano em que foi fundada a *Royal Statistical Society*, e para 1980, ano em que se criou em Portugal a primeira associação do mesmo género, e ainda para 1991, ano em que surgiu a actual SPE, precisamente mais de século e meio de distância. Quanta experiência perdida e quanto progresso adiado, ao longo de praticamente duas gerações? Nesta circunstâncias não é difícil aceitar que John Gower, estatístico inglês de reputada competência, terá dito, ao apresentar o balanço de uma visita que fez, na década de 80, a alguns departamentos em Portugal com actividade e interesse na área do delineamento experimental “... *most of the people are unaware of the last 40 years ...*”.

Afastando a compaixão podem fazer-se leituras convenientes desta situação aparentemente bizarra. Uma leitura pessoal consiste em olhar para aqueles que, dentro do atraso, se interessaram pela Estatística trabalhando no isolamento e de forma autodidacta, procurando saber aquilo que se tornaria inevitável saber. É justo que se guarde um pensamento de gratidão para todos eles, tanto os mais arrojados e brilhantes como os mais modestos, pois todos foram contribuindo para anunciar necessidades e preparar terrenos para a chegada de caminheiros mais habilitados e predispostos para conduzir a cruzada do ensino da estatística e da investigação nesta mesma área. O milagre aconteceu na década de 70 quando um grupo de jovens, entusiasmados com estudos e projectos emanados da atmosfera do então recente Centro de Estatística e Aplicações da Faculdade de Ciências de Lisboa, avançou para o doutoramento nesta área criando o embrião indispensável para o desenvolvimento e divulgação da Estatística que há muito se esperava (Branco, 1999). Daí até à criação da Sociedade Portuguesa de Estatística e Investigação Operacional (SPEIO), em 1980, foi um passo percorrido num instante.

A SPE dos nossos dias surgiu em 1991, quando as águas da Investigação Operacional e da Estatística foram separadas e os seus rumos ficaram devidamente individualizados. Daí até então a SPE cresceu, consolidou-se, e logo em 1993 deu início à realização de congressos anuais, uma actividade que exige grande esforço de organização mas que tem sido o motor de aglutinação de muitos estudiosos e utilizadores da estatística. A SPE é hoje uma sociedade com um grande número de sócios, quando a comparamos com outras sociedades de estatística de países com um volume de população próxima da de Portugal. Ao mesmo tempo que realiza o congresso anual faz publicar as actas do congresso, um esforço acrescido que traduz o empenho da SPE em manter este registo e fomentar a escrita dos resultados do labor dos participantes. Uma actividade que se tem vindo a realizar cada vez com mais profissionalismo e exigência. O autor destas notas desconhece que haja outra sociedade de estatística que faça congressos anuais com a publicação das respectivas actas. É neste contexto que se considera apropriado olhar para os congressos da SPE, em particular para as suas actas que anualmente se publicam, para se poder avaliar o volume de trabalho de organização que está em jogo e apreciar as contribuições científicas dos muitos participantes. O que poderemos concluir da análise destas contribuições?

3. Breve leitura das actas

Para apreciar a magnitude da participação nos congressos da SPE construiu-se a Tabela 1 cuja informação não contempla, intencionalmente, a contribuição, tanto em termos de comunicações como em artigos, dos convidados estrangeiros. Embora isto não seja relevante, há em média 3 convidados estrangeiros por congresso, pretende-se isolar a contribuição portuguesa. Num ou noutro caso poderão ter sido contabilizados participantes estrangeiros que não os convidados, mas como essa participação é rara consideramos que a sua inclusão, se ela existir, não vai perturbar as conclusões que se possam obter quanto à verdadeira contribuição portuguesa. Os números mostram que a partir do ano 2000 a maior parte dos congressos teve mais do que 200 participantes (excepções compreensíveis para 2001 (ilhas longínquas), para 2007 (Agosto e ISI em Lisboa) e para 2008 em que Vila Real chegou aos 195, quase 200). No global foram feitas 1585 comunicações (orais e poster) das quais 751, isto é 47%, vieram a ser publicadas nas actas. O gráfico da Figura 1 mostra de forma mais atraente e clara aspectos da informação contida na Tabela 1. Os dados da Tabela 1 foram obtidos a partir do livro de resumos dos congressos, que foi a fonte constante de informação para este exercício. A partir de 2003 o número de participantes deixou de ser incluído no livro de resumos, tendo a partir desta data sido usados os valores que aparecem no prefácio das actas. A informação que decorre da realização dos congressos nem sempre coincide com a que consta do livro de resumos e por isso é indispensável mencionar as fontes para esclarecer discrepâncias que possam surgir.

E o que estudam e investigam os estatísticos portugueses? Uma vez que a comunidade não é grande, poderá o seu interesse estar confinado a um reduzido número de temas? Não é isso o que a Tabela 2 mostra. Os títulos da última coluna foram seleccionados da longa lista de temas das comunicações livres dos vários congressos.

Tabela 1. Dados apurados nos vários congressos

Congresso	Participantes	Comunicações			Actas	
		Orais	Poster	Total	Artigos	Páginas
I-Vimeiro1993 14-17 Junho	134	42	14	56	41	570
II-Luso1994 20-23 Junho	121	43	8	51	32	483
III-Guimarães1995 26-28 Junho	176	56	21	77	53	699
IV-Funchal1996 14-17 Junho	120	47	26	73	46	429
V-Curia1997 11-14 Junho	162	34	41	75	50	523
VI-Tomar1998 09-12 Junho	121	23	35	58	32	414
VII-Ofir1999 13-16 Outubro	175	51	26	77	41	609
VIII-Peniche2000 04-07 Outubro	216	56	24	80	35	431
IX-P.Delgada2001 04-07 Novembro	107	41	24	65	29	437
X-Porto2002 25-28 Setembro	223	61	52	113	59	686
XI-Faro2003 24-27 Setembro	279	79	70	149	74	882
XII-Évora2004 29 Set.-02 Outubro	254	96	72	168	74	819
XIII-Ericeira2005 28 Set.-01 Outubro	220	79	44	123	61	716
XIV-Covilhã2006 27-30 Setembro	195	86	53	139	75	868
XV-Lisboa2007 19-21 Agosto	171	65	39	104	49	584
XVI-Vila Real2008 01-04 Outubro	210	98	79	177		
Total	2884	957	628	1585	751	9150
Média	180.25	59.81	39.25	99.06	50.07	610

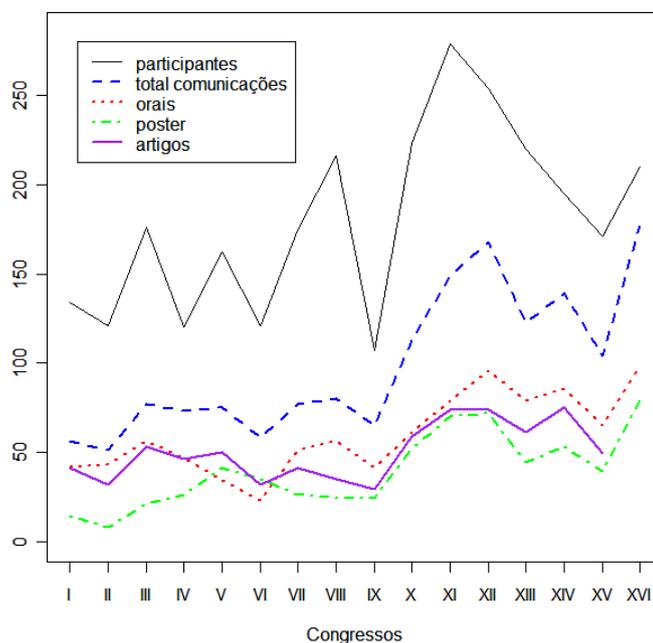


Figura 1. Representação gráfica dos dados dos congressos

Muitos dos temas, pela sua popularidade e interesse, aparecem repetidos em vários congressos, eventualmente com outros títulos. A lista revela que a diversidade de temas é grande e abarca não só aqueles temas que constituem o corpo central da estatística, como outros mais específicos, sem esquecer as aplicações. A sua intersecção com listas de áreas de investigação e interesse organizadas pela SPE e RSS, por exemplo, é grande.

Tabela 2 – Minicursos e uma lista de temas das comunicações livres

Congresso	Minicurso	Sessões de comunicações livres
Vimeiro		. Processos estocásticos e aplicações
Luso		. Estatística e aplicações
Guimarães		. Séries temporais
Funchal		. Ensino da Estatística
Curia		. Teoria de valores extremos e aplicações
Tomar	Tópicos de sondagens	. Estatística multivariada e análise de dados
Ofir	Controlo estatístico de qualidade	. Probabilidades e Estatística
Peniche	Modelos lineares generalizados - da teoria à prática	. Estatísticas oficiais e sondagens
P.Delgada	Inferência sobre localização e escala	. Estatística em Biologia e ciências médicas
Porto	Modelos heterocedásticos. Aplicações com o software Eviews	. Estatística em ciências sociais
Faro	Séries temporais. Modelações lineares e não lineares	. Inferência estatística e aplicações
Évora	Análise de clusters	. Modelos de regressão
Ericeira	Equações diferenciais estocásticas	. Probabilidades e processos estocáticos
Covilhã	Outliers em estatística	. Aplicações à medicina e epidemiologia
Lisboa	Introdução aos métodos robustos	. Caracterização e detecção de outliers
Vila Real	Análise de dados espaciais	. Estatística Bayesiana
		. Análise de sobrevivência
		. Controlo de qualidade
		. Modelos lineares e delineam. experimental
		. Amostragem
		. Dados categorizados
		. Estimção robusta
		. Estatística espacial
		. Modelos lineares generalizados
		. Análise classificatória
		. Amostragem e reamostragem
		. Estatística em ciências da saúde
		. Econometria
		. Estatística em genética
		. Modelos de equações estruturais
		. Estatística aplicada à economia e gestão
		. Estatística e sociedade
		. Regressão logística

No congresso realizado no Funchal foram apresentados 5 cursos nos dias 9 e 10 que antecederam os trabalhos do congresso, mas o primeiro Minicurso, com a estrutura que já se tornou habitual, surgiu no congresso realizado em Tomar, tendo esta realização ficado instituída desde então. Os 11 Minicursos já ministrados cobrem tópicos fundamentais da estatística. Em termos gerais pode então afirmar-se que, daquilo que ressalta dos congressos, os temas de estatística que interessam aos congressistas são muitos variados e o seu número tem vindo a aumentar com a introdução de novas áreas de investigação (Estatística em genética, que é uma área moderna e muito apetecida, aparece no XIV congresso, depois de uma primeira abordagem, Aplicações à genética, que havia acontecido dois anos antes no XII congresso).

A apreciação crítica da investigação revelada nos artigos contidos nas actas é uma tarefa bem complexa. Uma tarefa que requer muito tempo, cultura, talento e uma visão abrangente da actividade estatística. Uma análise assim tão impertinente fica adiada para quando estiverem reunidas tais condições. Desta vez olha-se apenas para um aspecto que interessa particularmente ao autor destas notas e que corresponde à convicção, que resulta da sua observação do que ouve e do que vê nos congressos, de que a comunidade estatística portuguesa tem vindo a interessar-se cada vez mais pela estatística aplicada, tomada no sentido da estatística que usa dados, reais ou artificiais (pretender distinguir entre estatística teórica e aplicada, num sentido mais lato, não é tarefa fácil, pois nem

sempre é claro onde acaba uma e começa a outra). Para analisar a consistência desta convicção foi efectuada uma contagem do número de artigos em que é feito o uso de dados reais e do número de artigos em que há recurso a dados simulados ou a dados hipotéticos. Todos os dados reais são contabilizados, tanto os que estão associados a um problema prático que o autor pretende resolver, como os que são copiados da literatura para serem usados na ilustração de métodos propostos pelo autor. Os dados artificiais são dados convenientes inventados pelo autor ou surgem como resultado de processos de simulação e são usados para efectuar comparações de vários métodos ou estimadores assim como para avaliar as suas sensibilidades e desempenhos.

Na Tabela 3 encontra-se o número de artigos publicados nas actas (Artigos), não incluindo, como na Tabela 1, os textos dos convidados estrangeiros, o número de artigos com dados reais (D. reais), o número de artigos com dados simulados (D. simulados) e o número total de artigos com dados (D. total). Note-se que este total é o número de artigos com dados, reais ou simulados, o que não coincide com a soma porque há artigos que usam os dois tipos de dados.

A Tabela 3 mostra o resultado dessa contagem, uma tarefa sedenta de tempo e paciência e que mesmo aparentemente satisfeita nos pode trair com casos menos claros. De facto casos há em que só nos apercebemos da existência de simulações quando mergulhamos profundamente no artigo, trabalho que na verdade não foi feito.

Analisando a Tabela 3, ou olhando para a Figura 2, onde se representam as proporções dos vários tipos de artigos considerados, não se vê razão forte para rejeitar a convicção inicial. Na verdade a proporção de artigos com dados, reais ou artificiais, é sempre maior do que a proporção dos outros artigos (onde não foram detectados dados), à excepção dos congressos II e IV, e a sua tendência não dá sinal de abrandamento.

Tabela 3. Artigos com dados reais e dados simulados

Congressos	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV
Artigos	41	32	53	46	50	32	41	35	29	59	74	74	61	75	49
D. reais	19	11	21	17	27	15	11	11	15	24	29	33	30	35	29
D. simulados	11	6	9	3	11	13	9	10	8	22	32	22	11	24	11
D. total	26	15	27	19	37	24	20	21	20	42	52	52	36	53	36

Note-se que a proporção de artigos com dados reais é muitas vezes maior do que a proporção de artigos onde não foram detectados dados, reais ou artificiais e é, com excepção do congresso XI, sempre maior do que a proporção de artigos com dados simulados. No entanto, apesar da importância dos dados reais, neste contexto, verifica-se, com insatisfação, que: i) muito raramente os dados reais ou parte deles, são incluídos no texto, prática que, infelizmente, é muito corrente; ii) muito raramente são dados associados a um projecto em que o autor do artigo esteja envolvido, sendo antes dados surgidos nas investigações de outros cientistas; iii) com frequência os dados reais aparecem repetidos em vários artigos do mesmo autor, situação que também não é estranha noutras áreas do trabalho científico.

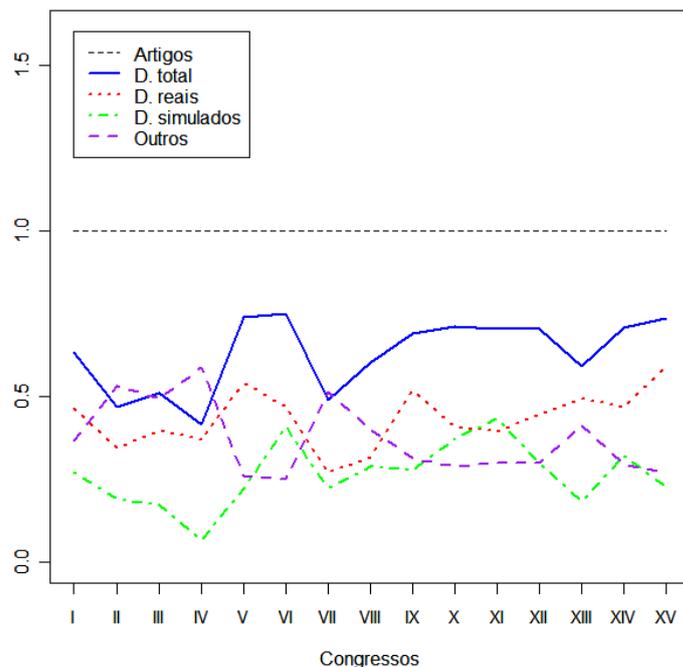


Figura 2. Gráfico das proporções dos vários tipos de artigos

4. Comentários finais

Em face da dimensão populacional do país, do interesse muito recente pelo ensino e investigação em estatística, que, em termos significativos, não vai além de 30 anos, e do escasso número de grupos de estatísticos que nas nossas universidades tem a responsabilidade do ensino avançado da disciplina, havendo apenas um departamento de estatística nas universidades portuguesas, pode dizer que a participação a que se assiste nos congressos da SPE é muito grande. A dedicação e a seriedade que a SPE tem colocado na organização dos congressos anuais e na publicação das suas actas são notáveis. A presença regular de investigadores estrangeiros convidados para participar em congressos ou visitar os grupos existentes com os quais mantêm colaboração, e a deslocação de investigadores portugueses a escolas estrangeiras para colaborar com os seus investigadores, são hoje acontecimentos frequentes e mostram que há um envolvimento com a comunidade internacional, com os novos problemas e com os novos métodos, o que só pode significar uma actualização permanente com o progresso da estatística. Muitos dos campos em que a disciplina se divide interessam aos investigadores portugueses, incluindo os campos mais recentes criados por exigência das novas tecnologias.

O movimento explosivo do passado pode estar agora numa fase de repouso em consequência das alterações, verificadas nos programas curriculares do ensino superior, motivadas pela introdução do Processo de Bolonha que aparentemente veio reduzir a já limitada intervenção da estatística no currículo. Não se conhecem ainda as consequências do novo sistema.

A pressão e exigências do mercado da produção e divulgação de trabalhos de investigação científica estão também a colocar interrogações que vão no sentido de questionar o modelo seguido no processo de publicação das actas da SPE, sobretudo o uso do Português e a menor divulgação internacional dos artigos. Neste momento a SPE está em negociações com vista a alterar o modelo de actas. Se essa alteração se vier a concretizar as actas passarão a ser publicadas numa nova “*International Book Series in Statistics*” da Springer o que implica que tenham de ser escritas em inglês e que o painel de avaliadores seja alargado de forma a incluir avaliadores de outros países.

A Estatística dos nossos dias está a ser chamada para problemas novos, gerados pelo desenvolvimento tecnológico. As preocupações do passado em conceber métodos para pequenas amostras deram lugar às preocupações do presente em saber como analisar grandes volumes de dados. Problemas de estatística caracterizados por grandes dificuldades de cálculo já não são mais problemas e a estatística experimental avança a passos largos alimentada pelas facilidades computacionais.

No passado a SPE deitou sementes numa espécie de deserto, de onde veio a florescer uma comunidade que proporcionou um desenvolvimento interessante da actividade estatística. Hoje, em face de conjunturas dos tempos modernos, o modelo de produção dessa actividade pode estar desadaptado da realidade vigente, mas conhecendo o sucesso da SPE ao semear no deserto é natural ter esperança e confiar em que a criatividade e perseverança da sua direcção e a fidelidade dos seus sócios chegarão para enfrentar a crise e continuar a conduzir a SPE no caminho do progresso.

5. Referências

Branco, J. (1999). Aspectos da realidade estatística em Portugal. *Boletim Informativo da Sociedade Portuguesa de Estatística*, **1**, 18.

Rosado, F. (2005) (ed.). *Memorial da Sociedade Portuguesa de Estatística*. Sociedade Portuguesa de Estatística, Lisboa.



Introduzindo *Outliers*

Fernando Rosado, *fernando.rosado@fc.ul.pt*

*Universidade de Lisboa
CEAUL e DEIO / FCUL*

quos fama obscura recondit¹

1. A Estatística como Ciência

A Ciência em geral e a Estatística em particular é uma nobre actividade, necessária ao corpo e ao espírito, indispensável ao bem-estar e à felicidade. Mas, a ciência é cara. Assim, só os ricos a podem praticar ... e os pobres se a praticam ficam mais pobres. Embora exigindo grande esforço e dedicação a solução deve estar em (apesar de tudo) fazer ciência para caminhar na saída daquele dilema. É o que se exige aos estatísticos portugueses congregados em torno de um projecto líder – a SPE! Esta, a Sociedade Portuguesa de Estatística reúne investigadores e fazedores de Ciência. No recente *Memorial*, Rosado (ed) (2005), diversos autores escreveram sobre a investigação. A actualidade dos capítulos mantém essa edição como bastante importante na consulta e na construção de uma história da investigação. Esta temática é essencial para a Ciência e para a Estatística.

A Estatística afecta a todos e atinge a vida em muitas situações. Como cidadãos ajudamos a fornecer informações estatísticas – a nossa própria entrada no mundo e a saída dele são registadas, para criar índices e taxas – e (através da publicidade) todos os dias nos tentam convencer de qualquer coisa ou mesmo enganar-nos à custa de factos e argumentos estatísticos.

É a importância do indivíduo / da observação.

A administração de uma comunidade, através das suas instituições de governo e comércio, depende muito das informações estatísticas e essa dependência aumenta à medida que o negócio intervém, cada vez mais, no planeamento da vida económica e social. Os propagandistas, gestores e dirigentes administrativos que utilizam (e que às vezes deturpam) as estatísticas são bastante numerosos e a eles podem juntar-se os mais diversos utilizadores – desde os estudantes de ciências sociais aos políticos. Todos empregam factos e métodos estatísticos para construir bases para políticas. Tais factos e métodos também têm um lugar muito importante no desenvolvimento da sociologia e economia como ciências, são relevantes para os experimentadores, na maior parte dos ramos da biologia e são fundamentais para aqueles que trabalham nas ciências mais exactas como a física, a química ou a engenharia.

É a importância da Estatística.

As ideias estatísticas estão pois, na base de muitas teorias e, de facto, uma "abordagem estatística" é talvez uma das facetas mais características da ciência moderna. Finalmente, a estatística como matéria é naturalmente do maior interesse para o grupo relativamente pequeno dos estatísticos profissionais.

¹ Meditando sobre a investigação. Expressão primorosa de Virgílio (Eneida, V, 302) glosada, entre muitos outros, por Santo Agostinho (*A Cidade de Deus*, volume I, Livro VII, Capítulo III, p. 611 e seguintes. Serviço de Educação. Fundação Calouste Gulbenkian. 1991). Na dicotomia entre a "razão menor" e uma "razão mais alta" deve o estatístico ter como objectivo (apenas) o conhecimento que lhe permite cobrir as suas necessidades científicas básicas? Em alternativa, esse deve ser um estádio inicial tendo por objecto a sabedoria estatística onde (ainda) admite a (enorme) importância dos "detalhes científicos" daqueles a quem uma obscura fama esconde – chamemos-lhes *outliers*; que são estimuladores da investigação e podem ser originados pelos valores discordantes de uma amostra – uma minoria. São esses "menores" que fazem avançar a ciência?! Neles está a força!

Como resultado das várias maneiras de encarar o assunto, a palavra estatística e as suas associadas, estatístico (adjectivo) e estatístico (substantivo), têm vários significados. Em primeiro lugar temos as definições dos dicionários, em que estatística se refere, no singular, ao assunto como um todo e, no plural, aos dados numéricos.

No senso comum – para o vulgar utilizador – as estatísticas são apenas números. Ele, tem tendência a pensar que um estatístico é principalmente uma pessoa que conta o número das coisas.

Para um economista, habituado às ideias qualitativas da teoria económica, estatístico é quase sinónimo de quantitativo. Para um físico, estatístico é o oposto de exacto visto que, para ele, a estatística é uma matéria que acima de tudo diz respeito a grupos e possibilidades mais do que a certezas. Para o cientista e investigador que está habituado a obter conhecimentos através da realização de experiências em condições que pode controlar, os métodos estatísticos são aqueles que se empregam quando é impraticável ou impossível um controlo experimental rigoroso. O campo de aplicação da estatística, na sua maior parte mas de modo nenhum totalmente, é económico e assim, o estatístico às vezes é considerado como uma espécie de economista. Por outro lado, como os métodos estatísticos são basicamente matemáticos, muitas pessoas pensam, ainda hoje, que o estatístico é uma espécie de matemático. Quase se poderia dizer que o matemático aceita o estatístico como economista e que o economista o considera um matemático. Alguns (poucos?!) pensam que os métodos estatísticos são tão pouco rigorosos que qualquer pessoa pode "provar" seja o que for; outros admitem que, demasiado severos, nada provam. No outro extremo situam-se aqueles que defendem que, como meio de aumentar os conhecimentos, o poder da estatística é ilimitado e quase mágico.

É habitual começar um livro, por exemplo sobre Estatística, definindo e ilustrando o assunto que vai abordar. Um tratado onde a estatística seja o tema principal, obviamente, não é excepção. Uma leitura (ao acaso) nas páginas iniciais de qualquer obra sugere duas perspectivas para as definições preliminares. Muitas vezes são breves e superficiais e outras inserem-se em campos particulares que limitam o próprio texto. Abordemos esta questão na sua maior generalidade. Ao encarar o tema da Estatística várias considerações se podem formular. E a primeira é a de considerar que é, ao mesmo tempo, uma ciência e uma arte. É uma ciência pelo facto de os seus métodos serem basicamente sistemáticos e terem aplicação geral e é uma arte porque o êxito da sua aplicação (também) pode depender da experiência e do engenho do estatístico e do seu conhecimento do campo onde actua. Contudo, não é necessário ser-se estatístico para apreciar os princípios gerais que lhe estão subjacentes. Como ciência, a Estatística e, em particular, os métodos estatísticos fazem parte do método científico em geral e baseiam-se nas mesmas ideias e processos.

Assim, a Estatística, tal como as outras disciplinas, está sempre em evolução. Ela é suportada por uma teoria... e, portanto, também e acima de tudo, é evolutiva! Uma teoria é um conjunto de princípios fundamentais de uma ciência ou de uma arte com formulação de uma doutrina acerca deles.

A Estatística é uma ciência porque, em síntese, desenvolve o conhecimento rigoroso e racional de um vasto ramo do saber e com as mais diversas aplicações. Por isso, deve ser um conjunto organizado de conhecimentos baseados em relações objectivas verificáveis e dotadas de valor universal.

É pacífico, vulgarmente aceite e em um bom dicionário pode encontrar-se que: uma definição para Estatística é uma ciência que tem por objecto o agrupamento metódico dos factos sociais que se prestam a uma avaliação numérica – da população, da natalidade e da mortalidade, dos rendimentos e dos impostos, das produções agrícolas, da criminalidade, da religião, etc. Numa perspectiva mais restritiva, por vezes também se elege a palavra Estatística para designar um ramo das matemáticas aplicadas que recorre ao cálculo das probabilidades para estabelecer hipóteses com base em acontecimentos reais, com o fim de fazer previsões. O avanço tem feito afirmar mais a primeira em detrimento da segunda.

A Estatística é a ciência dos dados, também aplicada porque a pesquisa, muitas vezes, visa também uma aplicação. A Estatística é interessante e útil porque fornece estratégias e instrumentos para trabalhar os dados de modo a melhor "entrar" em problemas reais. Dados são números (ou a falta deles) inseridos num determinado contexto ou experiência. Mas, determinar a média de 50 números é puro cálculo aritmético, não é Estatística. Discernir sobre aquele valor 50 e decidir se temos uma pequena ou grande amostra e, em cada caso, concluir sobre a discrepância de determinado valor (mesmo que usando a média atrás calculada!) já é Estatística. Embora a Estatística se possa considerar

como uma ciência matemática, ela não é um ramo da matemática e não deve ser ensinada como tal. Cada vez mais, podemos falar em pensamento estatístico que suporta e se apoia na teoria da decisão.

2. Ciência Estatística: Inferência e Decisão

A Estatística, na prática, exige julgamentos. É fácil listar as hipóteses matemáticas que justificam o uso de determinada metodologia, mas não é tão fácil decidir quando esse método pode ser "com segurança" utilizado na prática. A experiência torna-se aqui fundamental. Mesmo na vertente científica mais simples – e menos polémica? – podendo ser, como já vimos, a Estatística admitida como um ramo das matemáticas aplicadas, o objectivo final, na maior parte das vezes, está relacionado com predições. Assim encarada é a vertente prática da Estatística que se está a salientar. Mas, a predição está directamente ligada à Inferência e à Decisão. Toda a teoria ilumina a prática e esta informa a teoria, numa relação dialéctica. Sempre que se questionam os utentes (principalmente os de maior interesse prático), surgem "sugestões" que se desejam "com (muitos) estudos de casos" e "propostas de trabalho" com "menos teoria e mais prática". No entanto, o apoio teórico é sempre reconhecido e... deve estar sempre disponível e por perto!

Chegamos assim, à Teoria da Decisão Estatística. É nela que se fundamenta e onde (cada vez mais) está a génese da "profissão estatístico". Sobre este tema – teoria – respigamos algumas ideias basilares de Murteira (1988).

"Apesar de o homem ser chamado diariamente a tomar decisões só muito recentemente os problemas que estas suscitam começaram a ser tratados segundo uma óptica científica" (*ib*; p. 97). Historicamente, "a teoria da decisão estatística deve-se essencialmente a A. Wald que seguindo a tradição de Neyman-Pearson alargou consideravelmente os horizontes abertos por estes, tirando partido do desenvolvimento da teoria dos jogos realizada por von Neumann e Morgenstern. O grande mérito de Wald (...) consistiu na contribuição para uma argumentação onde (...) em termos gerais, os procedimentos clássicos são casos particulares da decisão estatística" (*ib*; p. 108-9). No entanto, para evitar confusões, importa que se clarifique que "(...) a teoria que vai desenvolver-se diz respeito à decisão individual e não de grupo. (...) a teoria que vai estudar-se não pretende substituir o decisor – mas, sim fornecer um conjunto de regras que auxiliem o decisor (...). Em termos gerais pode dizer-se que se está perante um problema de decisão quando se torna imperioso escolher ou optar entre, pelo menos, dois cursos de acção" (*ib*; p. 97). Mas, (haverá ainda quem considere que) a Estatística não é uma teoria? Poderemos admitir que se trata de "um instrumento ou alfaia cujas aplicações mais relevantes se situam, naturalmente, no domínio da investigação científica"?² "Nos problemas de inferência estatística ou de decisão estatística trabalha-se, quase sempre, no quadro de um modelo probabilístico ou, pelo menos, com uma forte componente probabilística" (*ib*. p. 23).

3. A necessidade de *outliers*³

Juntando esta componente, os modelos são muito importantes (fundamentais!) na investigação científica. Além disso, (também) na modelação surge de capital importância a amostra e cada um dos seus constituintes – (em particular) as observações e a sua dimensão. Na pesquisa de *outliers* numa amostra é, pois, de um problema de investigação que se trata... E, como tal, vai gerar (ou criar a necessidade de se apoiar em, pelo menos) uma Teoria! Alguns cientistas tomam Teoria como sinónimo de hipótese. Mas, a Teoria é distinta da hipótese e da ciência como sistema total. Porque, no método científico, a hipótese é uma fase anterior à Teoria e é parte integrante da ciência, teórica ou aplicada. A Teoria opõe-se à *praxis* ou acção mas são complementares. Como sabemos, o método científico atravessa várias fases: observação e experiências, hipótese e formulação da lei geral ou Teoria. Esta é pois uma hipótese já confirmada pela experiência e que faz parte integrante da Ciência. Mas há vários tipos de Teoria e de Ciências. No entanto, só são possíveis dois tipos basilares de Teoria: as dedutivas e as indutivas. Nas dedutivas, existe um conjunto de proposições válidas ou verdadeiras (teorema) que se constrói a partir de um grupo de proposições primitivas (axiomas) pela

² Murteira (1988), citando Gustavo de Castro, 1952, A Estatística Matemática uma Alfaia Científica. *Rev. Med. Veterinária* 52-64.

³ Breve síntese deste tema. O desenvolvimento pode ser seguido em Rosado (2006).

aplicação de certas regras de inferência. Nas indutivas temos um conjunto de proposições verdadeiras ou prováveis (teoremas, axiomas e definições) que se elabora a partir de vários casos particulares por um processo de inferência imediata e generalizante. Para muitos pensadores a conclusão da indução é apenas uma probabilidade. E, a "probabilidade de uma lei" cresce com o número de casos que a confirmam. Se, por diversos testes "confirmamos um *valor discordante*" temos o conjunto, isto é, o suporte das regras que podem formar – uma ou a – teoria dos *outliers*.

Uma das formas mais gerais de definir Estatística é aquela que a considera um método de proceder à escolha de alternativas de acção em face da incerteza, pelo recurso à recolha e interpretação de dados sobre os fenómenos em estudo. Assim, nesta generalidade, a Teoria dos *Outliers* em Dados Estatísticos torna-se uma mais valia para a Ciência Estatística. Na sua construção algumas dificuldades surgirão; contudo os obstáculos não parecem intransponíveis e é natural que se assista a grandes desenvolvimentos futuros. Evidentemente, os métodos de interpretação estatística, na sua concepção teórica, são rigorosos e conduzem a conclusões válidas do ponto de vista científico. Além disso, a própria qualidade dos dados utilizados, afectando as conclusões, não pode constituir base para a acusação desses métodos. A má qualidade pode pôr em causa a metodologia. A boa qualidade dos dados é um bem estatístico que se deseja. E, decerto, a qualidade melhora-se com a Teoria dos *Outliers*.

Em conclusão – uma ou a – teoria dos *outliers* deve fornecer um conjunto de regras que auxiliem o decisor além de construir instrumentos que permitam avaliar a qualidade da decisão. Com esse desiderato... devemos prosseguir! Mas, para crescer e se desenvolver, a teoria deve ser transmitida a outrem que (afortunadamente?) a expandirá com a sua contribuição!

4. A Fortuna⁴ / O Acaso decide!?

Admitamos uma análise de *outliers* em dados estatísticos. Dividamos – separemos – os dados em estudo em duas partes (dois e só dois grupos?): "os eleitos", que admitimos em maior número – a maior parte – e "os suspeitos". Estes (sempre presentes?!) estão em muito menor quantidade pois vulgarmente⁵ consideramos apenas um ou dois valores discordantes. Não há qualquer motivo assinalável para aquela escolha. No entanto, ela é (quase sempre) feita. A confirmação está na (cada vez maior) utilização dos testes de *outliers* nos mais diversos domínios científicos e pelos diferentes especialistas também nas aplicações, inclusive dos pacotes estatísticos. Todos desejam melhorar a qualidade dos seus trabalhos e conclusões através de uma "purificação dos dados". Mas, poderão os suspeitos conter mais e melhor informação que os eleitos? Qual é a causa que faz eleger estes em detrimento daqueles? Porque é que os suspeitos – a quem uma obscura fama esconde – não são os eleitos das análises estatísticas mais eloquentes? É que se trata de um assunto muito importante este de mostrar⁶ quais são "os verdadeiros"; embora possa não ser deles que nos vêm todas as (melhores!) pistas e, também, fazer ver todas as fragilidades das conclusões. Divididos os dados estatísticos entre eleitos e suspeitos, devemos questionar "quem é" ou "porque é" eleito. Quem atribui essa "condição estatística"?

O Acaso⁷ é a única coisa que não acontece por acaso. A Estatística⁸ é muito antiga mas tem uma história curta. Só entrou nas academias⁹ no segundo quartel do século XX e o principal arquitecto foi Fisher – justamente apelidado de fundador da moderna ciência estatística. Com certeza, a Fortuna é senhora; é ela que, mais conforme o seu capricho do que conforme a justiça, assegura a todos "a eleição" ou, pelo contrário, "a obscuridade".

Mas, quantos eleitos e quantos suspeitos? Se a Fortuna decide quem são os obscuros – aqueles que têm "a força" – porque não consegue essa honra para si própria – é caso para pensar? Porque sofreu uma

⁴ A deusa romana que, tal como a congénere grega Tyche, fazia a seu capricho, a felicidade e a infelicidade.

⁵ Também por razões científicas!

⁶ É sempre a procura da Verdade que está em causa!

⁷ A propósito de *outliers*, revisitemos este tema. A frase é de Almada Negreiros (Cf. p. 125 de *Matemática e Cultura*. Furtado Coelho *et al.* 1992. Edições Cosmos). Já foi tema de conferência de Tiago de Oliveira (*ib;* p. 125-49). A estatística casa bem com o acaso e ambos criam necessidade. É um tema recorrente que também já gerou o título para uma edição SPE – *Estatística com Acaso e Necessidade*; Actas do XI Congresso Anual.

⁸ Neste contexto, leia-se a "pequena excursão" apresentada por Tiago de Oliveira (*ib;* p. 125-8).

⁹ Sobre este assunto são importantes os artigos de Efron e Rao em Rao e Székely (2000).

fortuna adversa? Nesse caso ela, que nobilita os outros mas a si própria não pode nobilitar, é adversária de si própria! Os eleitos devem estar sempre em maioria? E, merecem mais atenção? Muitas vezes são escolhidos em razão da maior importância ou da contribuição para o estudo. E de entre os escolhidos, quais são "os mais fracos"? São todos igualmente bons? Qual é então a razão (ou a causa!) que obriga alguns eleitos a terem menor importância (por que estes devem existir!). E a comparação entre os "eleitos menores" e "os "suspeitos"? Estes, já chamados *outliers*, podem ter muito mais valor. É um *outlier* que abre o acesso a uma análise estatística mais profunda. Essa pode ser a origem de um trabalho de excelência. É a primazia em qualquer estudo de *outliers* que pode fazer a diferença entre um estatístico e um utilizador da estatística. Assim também na investigação. São os *outliers* – a quem uma obscura fama esconde – quem confere vida aos dados. Com certeza que o acaso é, quem mais assegura a cada estatístico a notoriedade ou a obscuridade. Não se julgue mais digno de honra, aquele que é agrupado nos eleitos. Mais fortes, são aqueles a quem o acaso – a mãe natureza – deu a condição de possuir muito mais informação (estatística). Para os eleger crie-se (pelo menos) uma teoria! A Fortuna / O Acaso também decide na investigação!?

5. *Outliers*: um caminho na investigação

A experiência pode tornar a investigação numa paixão. A investigação científica de base, tanto como a aplicada, constitui uma expressão significativa do domínio do homem sobre a natureza. Por sua vez, a investigação é *leitmotiv* na procura da Verdade. Esta pesquisa, por seu lado, estimula a grande reflexão (de sempre!) que deseja o concílio entre a razão e a fé que, muitas vezes, se questiona. Elas são as duas asas com as quais o espírito humano se eleva para a contemplação da verdade. Se a investigação científica proceder segundo métodos de rigor absoluto e se mantiver fiel ao seu objecto próprio não há oportunidade para qualquer dissensão. E, se a pesquisa é feita nos pequenos pilares – os de segunda linha – ainda mais fácil é a compreensão mútua dos “contendedores”; que não chegam a metas (supostamente) contraditórias.

Todas as grandes teorias começam com um pequeno passo e, muitas vezes, com base em dados (estatísticos). Pelo seu envolvimento em diversas matérias, a Estatística é transversal na Ciência e, por isso, a Investigação (em) Estatística é um grande pilar.

Na Investigação (em) Estatística, tal como noutra ciência, busca-se a verdade. A verdade? A procura é um caminho contínuo que se constrói com pequenos passos – indecisos e frágeis no princípio – mas, firmes quando a experiência e a solidez de conhecimento o permite. No entanto, “não há caminhos há que caminhar” (diz o poeta) pois “dos caminhos apenas ficam os rastos...” e, assim também na Estatística.

Caminhemos!

À medida que se caminha – mais aqui e menos ali – surgem perguntas laterais sobre “objectivos”, “interesse” ou “valor” da “ciência produzida”. No final está sempre a “procura da verdade” e qual o seu valor...

Decerto, as pequenas contribuições são sólidas porque facilmente “se controlam”, “assimilam” e “arrumam”. Assim, ocupam o devido lugar! As grandes, por sua vez, poderão ser mais frágeis pela vulnerabilidade nos pequenos suportes em que se apoiam. E são estes que, formando um todo, criam uma teoria.

Como na Estatística, nas diversas áreas do conhecimento também existem *outliers* – desde a física até à metafísica. Em todas elas, grandes interrogações se podem formular sobre o tema Investigação. Pesquisando *outliers*, onde quer que se encontrem, criamos Ciência – em particular, Estatística. Por consequência, abrimos (mais) um caminho na pesquisa da Verdade.

É a força dos menores!

Referências

- Murteira, B. J. (1988) – *Estatística: Inferência e Decisão*. Imprensa Nacional Casa da Moeda.
- Rao, C. R. e Székely, G. J. (2000) – *Statistics for the 21st Century. Methodologies for Applications of the Future*. Marcel Dekker, Inc. New York.
- Rosado, F. (Ed.) (2005) – *Memorial da Sociedade Portuguesa de Estatística*. Edições SPE.
- Rosado, F. (2006) – *Outliers em Dados Estatísticos*. Edições SPE.



Um olhar sobre a investigação estatística em Portugal

Carlos A. Braumann, *braumann@uevora.pt*

CIMA-U.E.

(Centro de Investigação em Matemática e Aplicações da Universidade de Évora)

Pede-me o Editor do Boletim uma opinião sobre a investigação estatística em Portugal. Através dos contributos informativos para este Boletim dos diversos centros de investigação onde a Probabilidade e a Estatística têm assento, o leitor ficará certamente com uma boa panorâmica do que decorre nos nossos dias nesta importante área de investigação, sem esquecer que há também investigação a decorrer fora desses centros, quer em instituições de ensino superior, quer em organismos públicos e empresas.

Não é pois disso que vos queria falar, mas antes de algumas impressões pessoais. Numa coisa certamente todos os que têm observado o pulsar da investigação estatística em Portugal estarão de acordo: houve um crescimento espectacular. E um barómetro é a participação nos 16 Congressos Anuais da Sociedade Portuguesa de Estatística (e, antes deles, nos Colóquios de Estatística e Investigação Operacional), que começaram com uma única sessão e que vão já em 5 sessões paralelas e várias sessões de posters, mantendo-se, com pequenas excepções que só confirmam a regra, uma elevada qualidade dos trabalhos apresentados. Estes Congressos, aliás, além de contribuírem para o convívio científico e pessoal entre os estatísticos portugueses, são também um factor de emulação para uma maior e melhor produção científica. Mas o que eu mais gostaria de salientar é o número crescente de jovens investigadores que eles revelam, o que é uma base sólida para um futuro auspicioso.

Mais ainda, a comunidade estatística portuguesa está a afirmar-se e a ser reconhecida a nível internacional, onde as publicações e comunicações com autores desta comunidade têm aumentado de forma notável. Não é por acaso que Portugal já tem uma revista científica internacional (REVSTAT) na área da estatística, que estatísticos portugueses ocupam lugares de algum destaque em sociedades científicas internacionais e que Portugal tem vindo a ser escolhido para a organização de muitas e importantes reuniões científicas internacionais, incluindo as de grande dimensão como o European Congress of Statisticians, a Sessão do International Statistical Institute e o COMPSTAT. E, deve-se dizer que, apesar de se ter a mania, verdadeira ou falsa, de que os portugueses não se sabem organizar, pelo menos tal seguramente se não aplica aos estatísticos, já que a qualidade da organização desses grandes eventos científicos tem sido vivamente elogiada.

Outra faceta que me parece muito relevante é a importância crescente que a investigação aplicada, virada para a resolução de problemas reais dos organismos e das empresas ou para o apoio à investigação nos mais variados domínios científicos, tem vindo a ter. É evidente que, pela sua própria natureza, a Estatística nunca esteve nem pode estar desligada das suas aplicações, que criam problemas específicos a requerer novas metodologias ou novos desenvolvimentos de metodologias existentes, os quais, por sua vez, permitem novas aplicações em áreas por vezes bem diversas das que os suscitaram. Mas o que eu queria dizer era que os trabalhos de aplicação directa a problemas reais aparecem hoje em maior proporção do que há umas décadas atrás, em que eram relativamente raros.

Isso deve-se não só ao envolvimento da comunidade académica e às suas preocupações de utilidade social, mas também ao facto de os organismos, as empresas e os investigadores de outras áreas científicas cada vez mais reconhecerem a importância do uso de metodologias estatísticas apropriadas e crescentemente recorrerem ao apoio da comunidade estatística. Aliás, o emprego de estatísticos em organismos e empresas tem crescido de forma notável e, neste momento, a produção de diplomados com formação estatística adequada não está a acompanhar as necessidades do mercado. Não é por

acaso que um estudo de 2007 do MCTES concluiu que o “índice de desemprego” (designação que eu dou ao rácio publicado nesse estudo entre o número de inscritos nos centros de emprego e o número de licenciados nos últimos 10 anos) dos licenciados em Matemática e Estatística é insignificante, o segundo mais baixo de entre os vários agrupamentos de licenciados (logo a seguir aos dos licenciados na área da segurança, onde, como sabemos, as admissões nos cursos são feitas em função das necessidades específicas do sector). Infelizmente, a sociedade instilou na juventude uma “alergia” à Matemática (e ao trabalho esforçado que o estudo da Matemática exige). Mesmo aqueles que não sofrem dessa “alergia”, desconhecem esta carência do mercado e, convencidos que o estudo da Matemática só habilita para profissões na área do ensino (onde os lugares estão saturados, situação temporária que irá rapidamente mudar), têm vindo a escolher outras licenciaturas. Podem estar a cometer um grande erro, até, porque, como se pode ver num recente artigo do colega Nuno Crato no Expresso, referindo-se a um estudo americano sobre as melhores carreiras profissionais, se verifica que os quatro primeiros lugares do *ranking* são ocupados pelas profissões da área da Matemática e Estatística.

Como consequência e, simultaneamente, como factor do progresso da investigação estatística, as exigências sobre a comunidade académica e a sua progressão na carreira têm vindo a crescer, o que se correlaciona positivamente com os níveis de publicação de trabalhos científicos. Felizmente, também têm vindo a melhorar ligeiramente as condições de trabalho desta comunidade, em boa parte graças aos centros de investigação e aos projectos de investigação, que têm conseguido colmatar os efeitos da preocupante degradação do financiamento das instituições de ensino superior e congregar os financiamentos indispensáveis aos meios informáticos, ao intercâmbio internacional, à participação dos estatísticos portugueses nas reuniões científicas internacionais e ao acesso às publicações mais relevantes. É bom que esses financiamentos acompanhem o progresso e o crescimento da comunidade estatística e que sejam igualmente acompanhados de financiamentos para a investigação aplicada e desenvolvimento em ligação com o tecido produtivo. Esperemos que, no concurso da FCT para projectos de investigação que está a decorrer não se assista à situação lamentável, observada em recentes concursos, de ver projectos de muito boa qualidade ficarem sem financiamento por limitações orçamentais, em boa medida decorrentes de se ter desviado parte das verbas de investigação para cooperação com instituições norte-americanas. Essas instituições são, sem dúvida, muito prestigiadas mas desconfio sempre de cooperações dispendiosas em que só uma das partes (a portuguesa) contribui financeiramente, beneficiando regamente a outra e assumindo uma posição de exagerada subserviência que o desenvolvimento científico do País já não justifica.

Naturalmente, a existência de um elevado número de frentes de investigação cobrindo as várias necessidades de desenvolvimento da ciência estatística começou por criar uma grande dispersão entre a relativamente pequena comunidade estatística portuguesa, grande parte dela doutorada fora do País em áreas muito variadas. Assim, só em algumas, raras, áreas estatísticas se fazia escola e se criavam massas críticas. Mas, à medida que esta comunidade foi crescendo, com o grande impulso que os programas de mestrado e doutoramento de muitas Universidades para tal deram, outras áreas estatísticas foram, em maior ou menor grau, ganhando massa crítica e visibilidade, para o que também ajudou a necessidade de organização dos centros e projectos de investigação. Há, porém, necessidade de consolidar essas áreas e de contribuir para que novas áreas se possam também erguer. Esse é talvez o principal desafio dos próximos tempos. Sinto-o claramente na minha própria área de investigação em equações diferenciais estocásticas e aplicações. Nesta área, graças ao esforço de alguns pioneiros e ao entusiasmo de muitos jovens, já se expandiu a teoria e se alargaram as aplicações a novas áreas de actividade (algumas das quais insuspeitadas, em que somos pioneiros a nível internacional e que criam necessidades específicas de desenvolvimento da própria teoria), começando a formar-se um corpo de investigadores com uma dimensão já razoável, mas que importa agora melhor organizar e consolidar.

Como sócio fundador da SPE, é com grande orgulho que constato o papel fundamental que esta sociedade e os seus sócios têm desempenhado no progresso extraordinário que se observou nas últimas décadas na investigação em probabilidade e estatística e nas suas aplicações. Cabe-nos a todos, no âmbito da SPE e das várias instâncias onde desenvolvemos a nossa actividade profissional e onde expomos o fruto dos nossos esforços, continuar este trabalho com exigência, clarividência e determinação.



Investigação (em) estatística: da biologia à estatística ou perdido no limbo

Tiago A. Marques, *tiago@mcs.st-and.ac.uk*

*University of St Andrews, Centre for Research into Ecological and Environmental Modelling,
Centro de Estatística e Aplicações da Universidade de Lisboa*

“Doing research is a creative process, there is no one way or right way of doing it; you need to discover what strategies work best for you.”

“Then one day, that random variable of all random variables, the mind, puts things together in a slightly different way and it is solved. The solution often then seems simple and obvious.”

Hamada & Sitter, 2004

Quando me foi pedido, pelo Professor Fernando Rosado, uma contribuição para o boletim da SPE, sobre a “Investigação (em) Estatística”, fiquei inicialmente um pouco surpreendido. Afinal, sou biólogo de formação e como biólogo me vejo. Depois de pensar um pouco mais sobre o assunto cheguei à conclusão que de facto a investigação que faço é actualmente mais sobre estatística do que sobre biologia, e como tal talvez pudesse escrever algo sobre o tema, partindo da minha experiência pessoal. Especialmente por ser baseado numa experiência ainda curta, aqui fica um aviso aos leitores: não prometo ser centrado, consistente, suficiente, não enviesado, preciso, ou ter qualquer outra propriedade de um bom estimador. Aliás, “variância e viés” poderia ser o outro título deste texto.

Assim que comecei a escrever estas linhas colocou-se me uma dúvida, bem ilustrativa em termos de estratégia de investigação. Deveria procurar ler sobre o assunto, e assim ver o que outros já tinham escrito sobre ele, ou tentar escrever algo sem influência directa de outras fontes? Neste caso, optei pela segunda abordagem, e escrevi uma primeira versão com base em “*nada*”. Foi interessante mais tarde fazer uma pesquisa e descobrir o trabalho “Statistical Research: Some Advice for Beginners” (Hamada & Sitter, 2004), onde várias das ideias que abordava eram também discutidas. E nele ler que “...*we find that it is easier to correct change, and work from an existing document than to start something new*”. Aí, já o embrião estava criado e pronto a ser moldado. Assim, tal como a investigação em geral, este texto foi um processo iterativo. E concordo plenamente com Thompson (2004) quando diz que Hamada & Sitter (2004) deveria ser leitura obrigatória para todos os estudantes que se iniciam numa carreira de investigação em estatística (tal como, acrescento eu, a discussão gerada no número seguinte do *The American Statistician*, e de que aqui e ali uso “pedaços”!). E também, como será evidente, com Stasny (2004), quando diz “*There is no one right way to do research. (...) Don’t worry if you don’t get it right the first time. Or even the 20th time. Keep trying, keep modifying, keep filling up the recycle bin.*”. Se temos um problema, há que olhar para ele até à exaustão, de vários ângulos e perspectivas, até perceber como abordar a questão. Por vezes também ajuda deixá-lo de lado uns dias, semanas se preciso for, e voltar a ele para o resolver de vez.

Para contextualizar as ideias que apresento interessa descrever um pouco do meu percurso como investigador. Depois de concluir o curso de Biologia, em 1998, na Faculdade de Ciências da Universidade de Lisboa, comecei logo a trabalhar num projecto sobre os impactos da barragem do Alqueva na conservação de peixes de água doce, na mesma equipa em que tinha feito o estágio profissionalizante.

Divagação I: Esse foi o meu primeiro erro, e aqui o explícito porque me parece frequente nos recém licenciados. Quando se acaba um curso superior e se pretende continuar a fazer investigação, devia ser feito um esforço para mudar de ambiente, de perspectiva. Parece-me um erro crasso, que parcialmente cometi, continuar na mesma equipa, no mesmo departamento, na mesma instituição. Mesmo que a equipa seja ótima, mudar é essencial, proporciona novas oportunidades, abre horizontes. Por vezes mostra que o que era ótimo só parecia ótimo, mas também mostra por vezes que o que pensávamos menos bom pode na realidade ser positivo. Não há ambiente só com defeitos, nem locais perfeitos, e a mudança ajuda-nos a apanhar o melhor de cada sítio por onde passamos, para quem sabe um dia no futuro o poderemos implementar num local nosso. Ficar “na mesma” é restritivo, pouco ambicioso, comodista, e diria que em nada contribui para o nosso futuro como investigadores. Mas é tipicamente português, a resistência à mudança e a empreendimentos de desfecho incerto (quem não se lembra do Velho do Restelo...?). Bem,... íamos na conservação dos peixes de água doce...

A vida do projecto consistia em, uma vez por mês, passar cerca de uma semana no campo a fazer pesca eléctrica numa dezenas de locais, e depois de volta a Lisboa introduzir todos os dados, preparar a saída do mês seguinte, escrever relatórios, realizar análises, etc. Na altura, apesar de haver objectivos gerais para o projecto, era pouco claro como estes seriam atingidos, e em particular como poderíamos recolher e analisar dados que permitissem atingir esses objectivos. Este facto criava em mim um certo desconforto, que me fez começar a olhar para estas questões, no fundo para a estatística e seus meandros, com bastante mais atenção. O contacto que mantinha com o Professor Dinis Pestana, que havia leccionado a cadeira de Biostatística no segundo ano do curso de Biologia, e que se manteve sempre disponível para ajudar na análise e interpretação de dados, levou-me a aceitar o desafio por ele lançado e a fazer o mestrado em Estatística.

*Divagação II: No meu entender, a vida é construída sobre acasos. Podemos ter muito talento, mas no final do dia a sorte (seja lá isso o que for), e todos os incontroláveis que ela controla, são preponderantes. E certamente concordariam comigo o genial Gamal Katif, que teria descoberto uma vacina para o cancro do pulmão em 2011. Ou o brilhante Simões Fonseca, que teria sido o inventor de uma fonte de energia virtualmente limpa em 2018. Concordariam... não fosse o primeiro ter morrido num ataque terrorista no Líbano quando ia para o liceu ou o segundo ter sido vítima de um condutor bêbado quando atravessava uma passadeira na Avenida de Ceuta. Mas certamente discordariam os excelentes actores e cantores que vejo na televisão a dizerem que só estão onde estão exclusivamente devido ao seu árduo trabalho e esforço. Para estar em qualquer lugar é preciso, antes de mais, alguma sorte. E afinal, como me mostrou o Nuno Sepúlveda, muitos outros pensam assim. Como Paul Auster, em *The Locked Room*: “In the end, each life is no more than the sum of contingent facts, a chronicle of chance intersections, of flukes, of random events that divulge nothing but their own lack of purpose”. E, apenas à posteriori se tornou óbvio, dois acasos condicionaram, bem ou mal, a minha vida: (1) ter feito o mestrado na FCUL, e não no ISA, porque o mestrado que pretendia originalmente não abriu nesse ano, e (2) um investigador natural da Costa Rica, com doutoramento tirado nos EUA, ter vindo para Portugal e pedido equivalência para o seu doutoramento, pelo que uma cópia deste estava no gabinete do meu orientador, que me passou e disse “vê lá o que achas disto!”. Eu vi, gostei do tema que ligava a biologia e a estatística na perfeição, e desde então nunca mais o deixei (já agora, Dinis e Russel, obrigado!).*

A minha tese de mestrado foi sobre amostragem por distâncias. Numa visita a St Andrews em 2001 para um *workshop* sobre este tema, na altura financiado, não pela fatal FCT, mas pela minha avó Gabriela, levou-me até ao Departamento de Matemática e Estatística da Universidade de St Andrews. Quase dois anos mais tarde iniciei lá a minha tese de doutoramento, sempre à volta do mesmo assunto, a estimação da abundância de populações animais. O doutoramento foi feito no CREEM (Centre for Research into Ecological and Environmental Modelling), uma unidade de investigação com poucas dezenas de membros, quase todos estatísticos ou biólogos, que partilham um observatório astronómico reconvertido. Desde o ano passado que me dedico, no âmbito de um pós-doutoramento, a estimar a abundância de cetáceos com base na informação dos sons que estes produzem e que são registados por

hidrofonos, no âmbito de um projecto multidisciplinar que inclui biólogos, estatísticos e engenheiros acústicos. No entanto, apesar de trabalhar no/para o Reino Unido, continuo a residir em Portugal, uma das vantagens do desenvolvimento de coisas como e-mail, internet, skype, net-meeting, desktop sharing, e outros chavões.

E neste processo todo, o que aprendi sobre a investigação estatística? Aprendi, em primeiro lugar, que não aprendi nada que consiga ser generalizado com facilidade, pois foi um percurso sinuoso e muito particular. O que resulta para umas pessoas não resulta necessariamente para outras, pelo que posso apenas fazer algumas considerações pessoais sobre o meu percurso como investigador, tentando retirar desse exercício ideias que possam ser úteis para a discussão geral sobre investigação estatística.

Penso que a investigação requer uma forma de pensar, uma formatação distinta, curiosidade, uma disposição natural e intrínseca para descobrir coisas novas. Ou seja, nem todas as pessoas seriam bons investigadores, da mesma forma que nem todas seriam bons médicos, marceneiros, advogados ou pintores (eu certamente seria péssimo a qualquer uma destas opções!). Apesar de essa forma de pensar se treinar, creio que muito é também inato. Isto não tem nenhuma base concreta, é apenas uma observação pessoal. A investigação é feita em grande parte por pessoas que também tem a responsabilidade de leccionar, nomeadamente os professores universitários. E em particular, um bom professor não faz necessariamente um bom investigador, nem vice-versa. Aliás, já conheci alguns excelentes investigadores que são sofríveis professores, e vice-versa. E assim sendo temos uma situação generalizada em que as mesmas pessoas tem que despende um tremendo esforço em duas áreas distintas, e para as quais as valências não são necessariamente as mesmas. Na melhor das hipóteses, podemos esperar que apenas uma das áreas seja descurada. Embora, não possamos esquecer, a interacção com os alunos possa também servir de estímulo, incentivo e inspiração à investigação. Assim sendo, eu diria a alguém que está a ler estas linhas sem saber o que fazer a seguir, que se o que gostaria de fazer na vida era investigação, não se tornasse professor. Mas, perguntam-me então, vai procurar emprego onde... pois, também não sei bem! Mas deviam de facto haver pessoas para leccionar, e outras para investigar. Os iluminados, que os há (aposto que muitos entre os leitores!), que fizessem as duas coisas.

Divagação III: Claro que esta é uma visão muito pessoal, provavelmente baseada num lento desempenho, insegurança e necessidade de preparação à exaustão, mas eu, cada vez que tenho de dar aulas, fico com a investigação num estado de latência. O ano passado, estive em St Andrews para assegurar uma cadeira em regime intensivo (estatística, 1º ano, 3 semanas: 1 hora por dia, e 1 prática e uma teórico-prática por semana). Durante esse tempo, posso afirmar que não fiz nada em termos de investigação, havia sempre mais um exemplo interessante para incluir na aula seguinte, uns slides adicionais para mostrar, um aluno a aparecer com dúvidas, e-mails de alunos para responder, trabalhos para corrigir, etc. Por isso admiro quem consegue fazer as duas coisas, com alguma qualidade em ambas, e frequentemente com cargas lectivas bem superiores. E pronto, acho que assim hipotequei em definitivo a minha hipótese de vir a ter um emprego numa universidade em Portugal.

A investigação estatística, e penso que a investigação científica em geral, requerem antes de tudo um ambiente apropriado. As condições no CREEM foram, e são, desse ponto de vista, excepcionais. Um espaço relativamente pequeno, mas não enclausurante, com várias áreas comuns (como uma sala de aulas/seminários, uma sala com uma vintena de computadores, uma área aberta, por exemplo para reuniões informais, uma zona para cafés, uma sala de reuniões) que facilitam a interacção entre investigadores. O grupo de pessoas que lá trabalha é diverso, com uma mistura de formações, culturas e nacionalidades, mas quase todos com um percurso que mistura a biologia e a estatística. A percentagem de pessoas que trabalha de porta fechada (literalmente) é reduzida. Todos os dias, por volta das 11:00, há uma pausa para o café (a famosa *coffee break*), em que todos participam, excepto se o trabalho estiver a apertar demasiado. Quase todas as semanas há um seminário, por alguém da casa ou convidado. É frequente haver reuniões de investigação para discutir um assunto específico de interesse comum, ou discussões de artigos científicos. As visitas de investigadores externos, ou pessoas e entidades à procura de consultoria para os seus problemas, são constantes. E três a quatro

vezes por ano, há *workshops* com participantes de todo o mundo, em diferentes tópicos, mas sempre à volta da análise de dados ecológicos e estimação de abundância de populações animais. Assim sendo, é quase impossível que não surjam ideias e projectos novos, e o difícil raramente é arranjar dados ou problemas interessantes para investigar mas ter tempo para responder a todas as solicitações que vão surgindo. As coisas de que mais sinto falta, por viver em Portugal e, no entanto, trabalhar no/para o Reino Unido, são todas as interacções informais que, no final do dia (ou de períodos bem mais longos!) fazem a diferença entre um ambiente altamente produtivo e um qualquer outro local de trabalho. Infelizmente, penso que em Portugal os ambientes são demasiadamente formais, as portas estão com demasiada frequência fechadas, não existem *coffee breaks*, a partilha de informação é frequentemente condicionada por jogos que nada tem a ver com a investigação, e na estatística em particular, parece que por vezes a investigação é feita sem grande interesse pela utilização que vai ser dada a essa investigação.

Certamente uma boa proporção dos leitores (num total de 3 a 4 que ainda estavam a perder o seu tempo nisto) se sentem neste momento injustiçados, e alguns provavelmente com razão (é o preço a pagar por qualquer generalização!), mas a verdade é que é muito fácil darmos por nós a trabalhar e resolver problemas que terão pouca ou nenhuma aplicação (não que ache que toda a investigação deva ser necessariamente aplicada, mas voltarei a esta polémica mais abaixo). Por exemplo, lembro-me pelo menos de um artigo que publiquei à cinco anos e que apenas foi citado por outros para referir que os problemas foram abordados, mas nunca citado por alguém que tenha implementado a metodologia sugerida num caso real. E no entanto, o problema abordado é real e existe. Mas isto também mostra a dificuldade que há em transmitir àqueles que aplicam as metodologias na prática os desenvolvimentos teóricos que acompanham essas mesmas metodologias.

Outra das coisas boas do CREEM para um jovem investigador na área da estimação de abundância de populações animais é que foram os seus investigadores a escrever os livros de referência na área. E se isso pode ser inibidor ao princípio, depois torna-se uma vantagem tremenda, uma fonte de informação privilegiada e disponível em tempo real. Assim sendo, eu diria que antes de começar a fazer investigação numa determinada área, deveria ser feito um trabalho prévio de pesquisa, para procurar saber, internacionalmente, quais os grupos mais activos na área, e depois fazer um contacto directo com as pessoas nesse grupo e tentar estabelecer colaborações. Mesmo em termos de bolsas da FCT para doutoramentos este facto é facilitado. Porque não pedir uma bolsa mista, com ligação a uma instituição estrangeira de qualidade internacional, em vez de ficar apenas numa instituição em Portugal (mesmo que essa instituição seja por si mesma de qualidade internacional, que também as temos!). É a mistura de ideias que traz vantagens.

Divagação IV: Sobre o inibidor que pode ser trabalhar com nomes sonantes.... Quando em 2001 fui ao workshop que acima referi, em St Andrews, aquele que mais tarde viria a ser meu orientador, Steve Buckland, convidou-me para almoçar, convidando também o Jeff Laake, um outro grande nome na área em questão. Nesse almoço perguntou-me se gostaria de voltar para fazer um doutoramento. Eu fiquei encolhido e disse que era difícil, balbuciando que não era fácil vir trabalhar com aqueles que escreveram a “bíblia” (referia-me à obra de referência em amostragem por distâncias, Buckland et al., 2001, saída nesse ano). O Jeff, percebendo o meu incómodo, disse, algo como: “Sabes, nós somos todos homens como os outros, vestimos as calças da mesma maneira!”; fez uma pausa de 3 segundos e disse: “Bem, de facto, acho que o Ken (Burnham) já as veste ao contrário de vez em quando!”). E isso quebrou o gelo.

Voltando atrás, uma outra particularidade do CREEM é que o principal software usado na análise de dados de amostragem por distâncias, o DISTANCE (Thomas et al., 2006) é desenvolvido pelo nosso grupo. Por causa disso (e também por causa dos *workshops*, que já acima referi, em que ensinamos os métodos bem como a usar o software), recebemos colectivamente centenas de e-mails por ano com perguntas práticas sobre os métodos. O manancial de informação, sobre casos de aplicação prática e em especial sobre os problemas que os investigadores envolvidos nestes enfrentam, levantam questões que não raramente acabam por resultar em colaborações profícuas. Além de métodos para os quais

existe software disponível para a implementação acabam por ser mais utilizados, dando visibilidade (e citações) aos seus autores. Qual é a lição a retirar? Talvez que o investimento em criar software para implementação de metodologias estatísticas tenha um retorno, e não deveria ser descurado. Mais uma vez contra mim falo, que nunca tornei público nenhum do código base das metodologias que desenvolvi no doutoramento! Claro que disponibilizá-lo-ei se for pedido... mas porque não me dei ao trabalho de, por exemplo, implementar um pacote de R para tal? Bem, a verdade é que acho que há a possibilidade do código “morrer” assim que correr noutros dados que não os meus...!

Como já se deve ter tornado óbvio, excepto para os mais desatentos, na minha opinião, o combustível ideal para a investigação são os problemas práticos. Ficar metido dentro do gabinete à espera que uma inspiração surja do nada é um processo bastante mais lento do que a interacção com os que têm problemas no terreno com a aplicação das metodologias. Isto não quer de forma alguma dizer que não devemos pensar mais abrangentemente do que para resolver o nosso problema específico. Parafraseando Donald B. Rubin, por sua vez referindo-se a uma intervenção de George Cobb, “*Work on a specific problem, but think abstractly.*” (Rubin, 2004).

Várias ferramentas me parecem úteis na investigação estatística, mas uma das minhas favoritas é certamente a simulação. Temos um problema, vamos ver como se comportam os estimadores usuais. Simulamos a violação de alguns pressupostos. Consideramos tamanhos de amostras realistas, em vez de usar propriedades assintóticas justificadas por tamanhos amostrais geralmente proibitivos, e vemos como se comportam os estimadores. Temos um novo estimador? Como reage em comparação com os métodos tradicionais. Assim sendo, parece-me que o domínio de uma linguagem de programação é fundamental para apoiar o processo de investigação. Neste contexto, não posso deixar de notar o contributo decisivo que ter aprendido a trabalhar em R (R Development Core Team, 2008) trouxe à minha investigação quando passei do mestrado para o doutoramento. Actualmente, é raro o dia em que não uso o R.

Divagação V: Quando fiz o mestrado apresentei um conjunto de, creio eu, 2400 simulações, em que gerei dados em Excel e corri o software Distance (Thomas et al., 2006,) pelo menos essas 2400 vezes, à mão. Se não me falha a memória, a geração de cada conjunto de dados levava uns bons minutos, depois tinha de importar tudo para o Distance, correr o programa, e era ainda preciso copiar à mão os resultados para os processar, etc. No total despendi três meses a fazer estas simulações. Quando comecei o doutoramento, um dos primeiros projectos a que me propus foi criar em Visual Basic um pequeno programa, que corria as 2400 simulações sobre os dados já criados em menos de uma hora. E depois uma forma de gerar os dados em R, chamar o Distance e ler os resultados directamente... aí passei de três meses de trabalho para três minutos! Perdem-se umas, ganham-se outras!

Hoje em dia, o sistema é tal que parece que a investigação tem algum mérito se for publicada numa revista de impacto internacional. Parece-me uma visão redutora, mas por outro lado, se a investigação é de facto de qualidade, deveria ser feito um esforço adicional para a sua publicação. Além de que a avaliação dos cientistas se faz cada vez mais cegamente à custa do número de artigos publicados (sendo que muitas vezes a qualidade e consistência destes é posta em segundo plano) e do factor de impacto das respectivas revistas. O que não deixa de ser um facto é que o processo de publicação, e todo o processo de transmissão de informação a outros, é uma parte essencial da investigação. Concordo com as palavras de Stasny (2004), “*No matter what path you are taking on your research, there are three things I advise every student to do: write, talk, and listen.*”, mas com a extensão óbvia a investigadores não estudantes. (1) *Write*: O facto de termos de pensar na melhor forma de transmitir as ideias a outros faz com que sejamos obrigados a reestruturar os nossos pensamentos, ver os problemas de ângulos diferentes, novas perspectivas, etc. (2) *Talk*: Com frequência, quando explico a alguém uma dúvida, ou coloco uma pergunta mais elaborada, ou escrevo um e-mail a solicitar explicações, acabo por encontrar a resposta à pergunta ainda antes do meu interlocutor a dar. Apenas a ordenação das ideias para expressar a pergunta foi suficiente. (3) *Listen*: ouvir outras pessoas falar, mesmo por vezes sobre áreas que nada tem a ver com a minha, abre a porta a inúmeras ideias, novos problemas e soluções alternativas para os existentes. Além de que promove uma cultura geral

científica, especialmente importante porque quando temos apenas um martelo na caixa das ferramentas, todos os problemas que vemos parecem um prego. Na minha opinião, os investigadores deveriam ir ao máximo de seminários que conseguirem, e em particular os alunos de doutoramento, deviam ser obrigados a tal!

Por outro lado, a investigação científica em geral requer dinheiro. Sem financiamento não há investigação. E vivendo nós num mundo de recursos limitados, isso faz com que um determinado campo se tenha de tornar apelativo para poder atrair dinheiro. Pode não ser bonito, mas é o mundo em que vivemos. Penso que ainda há um longo caminho a percorrer nessa área no que diz respeito à estatística. São necessárias mais aplicações em que a incorporação de estatísticos (e suas caixas de ferramentas!) numa fase inicial do processo tenha gerado claras mais valias, e depois publicitar esse facto. No estrangeiro os gabinetes de imprensa das universidades esforçam-se por divulgar os feitos dos cientistas das suas instituições, procurando activamente dentro destas tais casos de sucesso. Por cá, vejo esses gabinetes de imprensa (quando existem!) mais preocupados em mandar (inúteis?) comunicados sobre as posições do vice-presidente do conselho directivo sobre as propinas de doutoramento ou coisa que o valha. Não quero com isto dizer que um investigador se deva dedicar ao marketing e publicidade, mas a longo prazo parece-me que quem não conseguir vender a sua investigação terá problemas. Até porque, além de atrair dinheiro, é também preciso atrair pessoas, a massa crítica para o desenvolvimento de qualquer área.

Divagação VI: Para atrair alunos para a estatística ajudaria por exemplo divulgar os resultados de uma recente seriação divulgada pelo Dr. Nuno Crato no Expresso (ver original em http://www.careercast.com/jobs/content/JobsRated_Methodology), segundo a qual “Estatístico” seria a 3ª melhor profissão do mundo (nos EUA, mas pronto, coitados, eles também tem a sua Baseball World Series só deles!). Claro que uma análise mais aprofundada do tema revela problemas com esta (e possivelmente com qualquer!) seriação. “Matemático” é aparentemente a melhor profissão do mundo. Um erro, certamente! (isto era apenas uma provocação gratuita aos meus amigos matemáticos). Mas sinto-me bem: tenho de certa forma 2 das 5 profissões melhores do mundo: “Biólogo” aparece em 4º lugar. Não devem ter usado dados de Portugal! Mas falava eu de dinheiro...

Qual a investigação que deve ser financiada é uma pergunta relevante, e não apenas na área da estatística. A investigação científica em geral pode ser catalogada de acordo com diferentes sistemas. Um desses sistemas dicotomiza a investigação em fundamental (por vezes também apelidada de “blue skies research”) e aplicada. O balanço entre estes dois tipos de investigação é objecto frequente de debate, havendo desde aqueles que pomposamente se acham tão mais importantes quanto mais conceptual e exotérica é a área em que trabalham, até aos que não aceitam que se gaste dinheiro a estudar problemas que, na melhor das hipóteses, “não interessam nem ao menino Jesus”. Como sempre, a adopção de visões extremas é criticada excepto pelos extremistas propriamente ditos, mas na realidade encontrar o balanço entre as duas vertentes não é tarefa fácil. A questão é ainda mais complicada porque me parece natural pensar que, para cada país o equilíbrio entre investigação aplicada e fundamental deve ser dependente da realidade desse país. No extremo, inexistente mas útil como exemplo, um país sem problemas não precisaria de investigação aplicada. À medida que o número de problemas práticos num determinado país aumenta, maior deveria ser o financiamento atribuído à investigação aplicada. Assim sendo, parece natural que num país com muitos problemas práticos numa determinada área, o dinheiro destinado a investigação nessa área deveria ser preferencialmente canalizado para investigação aplicada. Podem criticar esta visão dizendo que estou a sugerir que em Portugal se corte no financiamento da ciência fundamental. Nada disso. Estou apenas a sugerir que provavelmente deveria ser feito um esforço sério e concertado para definir que áreas devem ser financiadas, e como.

Divagação VII: Num comentário mais geral, dá-me ideia que o estado descuro no sistema actual algumas das suas responsabilidades, colocando nos ombros de investigadores a responsabilidade de encontrar financiamento para responder a questões que um estado de direito deveria assegurar. Por

exemplo: ouvi há dias num debate um médico referir que havia duas correntes, uma que dizia que devia ser o estado a assumir as custas para implementar coisas como inquéritos nacionais de saúde (e que a mim me parece ser óbvio que assim seja!), outra que dizia que os investigadores deveriam arranjar dinheiro de forma independente para inquéritos que achassem relevantes. E no final esse financiamento acaba por ser obtido muitas vezes por intermédio do estado. É irónico e perverso: um investigador, da função pública e assim pago pelo estado, gasta o seu tempo (e assim o nosso dinheiro!) a inventar projectos que depois são financiados pelo estado, para responder a perguntas do interesse do estado e que este devia financiar directamente. Não sei se o comentário que ouvi corresponde à realidade ou não, mas já imagino um dia desses um comandante de polícia a fazer projectos para a FCT, pedindo 1 carro, 2 fardas, cacetetes, munições e 30 litros de gasolina para ir prender uma quadrilha qualquer a um qualquer bairro problemático (assumo, roubei esta ideia a alguém, mas falha-me a memória a quem... mas é sempre assim, as ideias raramente são verdadeiramente originais.).

Assim, deambulando chego a uma ideia que me parece consensual, a de que deve haver uma maior e melhor comunicação entre os investigadores da área da estatística e os de outras áreas científicas, pois da interacção entre as duas áreas resultarão certamente problemas e soluções interessantes. Esta preocupação está patente por exemplo nos recentes, e na minha opinião altamente louváveis, seminários do CEAUL em que pessoas de áreas não estatísticas foram convidadas para apresentar os seus trabalhos e problemas, como forma de lançar as sementes para um diálogo potencialmente proveitoso entre as partes. Para esse diálogo ser efectivo há também que primeiro, haver uma linguagem comum, e um esforço de ambas as partes para se entenderem. Não basta criticar a outra parte por não se esforçar.

Depois de tantas divagações resta-me concluir.

Neste momento trabalho num limbo entre a Estatística e a Biologia, tal como o Nuno Sepúlveda. Com o tempo, vou-me tornar cada vez mais um sofrível biólogo, e temo que nunca, por mais que me esforce, serei um estatístico. Mas sem dúvida, sinto-me cada vez mais capacitado em fazer a ligação entre as duas áreas. A investigação estatística apenas tem a ganhar da cooperação com a Biologia, ou com vários outros domínios científicos (e.g., Medicina, Física ou Ciências Sociais). É necessário, porém, que os estatísticos estejam dispostos a "sair da casca", e dar o salto para o outro lado, e que os biólogos (ou outros cientistas) tenham o discernimento, e a vontade, de procurar a colaboração numa fase inicial do seu trabalho. Mãos à obra!

Agradecimentos

Ao Nuno Sepúlveda pela troca de ideias ("and apparently ours do too...", LOL), pela leitura e comentários. À Luisa Chaves, que leu, comentou e assim melhorou o texto com a paciência do costume. Aos meus orientadores de mestrado e doutoramento, Dinis Pestana e Steve Buckland, que me iniciaram no processo de investigação. Esse é também um factor determinante para uma carreira de investigação, os mentores que temos. Eu, tive sorte.

Referências

- Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., & Thomas, L. (2001). *Introduction to distance sampling - Estimating abundance of biological populations*. Oxford University Press, Oxford.
- Hamada & Sitter (2004). Statistical research: some advice for beginners. *The American Statistician*, **58**: 93-101.
- R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, <http://www.R-project.org>.

Rubin, D. B. (2004). On advice for beginners in statistical research. *The American Statistician*, **58**: 196-197.

Thomas, L., Laake, J.L., Strindberg, S., Marques, F.F.C., Buckland, S.T., Borchers, D.L., Anderson, D.R., Burnham, K.P., Hedley, S.L., Pollard, J.H., Bishop, J.R.B. & Marques, T.A. 2006. Distance 5.0. Release 2. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. <http://www.ruwpa.st-and.ac.uk/distance/>

Thompson, E. (2004). The importance of “Why?”. *The American Statistician*, **58**: 198.

Stasny, E. A. (2004). Some random thoughts on doing research. *The American Statistician*, **58**: 200-202.



Surfando pelos mares da Biologia: peripécias de um jovem estatístico aspirante a biólogo

Nuno Sepúlveda, *nunosep@igc.gulbenkian.pt*

*Instituto Gulbenkian de Ciência, Oeiras
Escola Superior de Saúde Egas Moniz, Monte da Caparica
Centro de Estatística e Aplicações da Universidade de Lisboa*

Foi com enorme satisfação, mas também com alguma perplexidade, que recebi o inesperado convite do Prof. Fernando Rosado, para escrever algo sobre um tema tão vasto, como é o da investigação (em) estatística. Sempre pensei que estes convites fossem endereçados a investigadores seniores, que não sou, cujos muitos anos de experiência permitem ter uma visão mais esclarecida sobre essa arte/ciência que é fazer estatística. Para além disso, o meu percurso académico nos últimos 4 anos tem-me feito afastar, mas não na totalidade, dos domínios tradicionalmente de Estatística. É certo que os meus estudos iniciais contemplam uma licenciatura e um mestrado em Matemática Aplicada, perfil em Estatística e Aplicações, no Instituto Superior Técnico (IST), mas a minha investigação actual incide no campo da Biologia Quantitativa, estando a terminar a escrita da tese de doutoramento sobre assuntos tão estranhos para um estatístico, como o papel das famigeradas células T reguladoras no sistema imunitário.

Apesar de tudo, tenho feito um esforço para combater esse afastamento, tentando participar com alguma frequência em congressos da área, com a submissão de trabalhos de índole aplicada, mas sempre com a Biologia como pano de fundo. Curiosamente, ao escrever isto, sinto que esses trabalhos, que embora não fizessem parte da agenda principal do meu doutoramento, foram, de certo modo, responsáveis pelo desafio que me foi aqui proposto.

“O que escrever?” era a pergunta que punha a mim mesmo, quando pensava em começar a redigir este artigo. Após umas quantas trocas de emails com o Prof. Fernando Rosado, “consegui” perceber o que se esperava de um novato como eu: dar uma visão pessoal sobre a investigação (em) estatística. Optei, assim, por contar algumas das minhas vivências aqui em Oeiras, no Instituto Gulbenkian de Ciência (IGC). Pareceu-me ser a forma mais adequada de atingir os objectivos pretendidos com este texto.

Vamos lá então “surf” um pouco pela minha experiência estatística no IGC.

Um estranho numa terra estranha

Todas as histórias têm um início, e a minha não é excepção. Em Outubro de 2001, quis o acaso que fosse entregar um livro ao Prof. Daniel Paulino no quinto piso do edifício de pós-graduação do IST. No átrio dos elevadores, vi afixado num placar um anúncio de uma bolsa de investigação científica, que requeria os serviços de um matemático, ou estatístico, para um projecto sobre a patogénese do vírus de imunodeficiência humana (VIH). O trabalho seria realizado no grupo de Imunologia Teórica do IGC, sob a supervisão do Dr. Jorge Carneiro. Achei a temática do projecto bastante apelativa e intrigante, pelo que me candidatei de imediato à bolsa. A entrevista correu bem, e fui felizmente seleccionado. Senão, não estaria certamente a contar esta história.

Comecei a trabalhar no IGC no mês seguinte. O meu trabalho consistia em compreender a resposta das células T do sistema imunitário de um hospedeiro à infecção pelo VIH através de sistemas dinâmicos. Para isso tive de estudar, com afinco, alguns conceitos básicos de Imunologia, assim como aprender a construir e analisar sistemas dinâmicos. A parte estatística do trabalho seria analisar dados a recolher posteriormente em pacientes do Hospital de Santa Maria, de forma a calibrar o modelo e estudar o seu comportamento.

Nessa altura o Instituto era constituído por cerca de meia centena de investigadores/estudantes, sendo a maior parte deles “experimentalista”, ou seja, trabalhavam nos *wet labs*, onde se realizavam estranhas experiências *in vitro*, *in vivo*, ou mesmo *ex-vivo*. Os seus interesses científicos centravam-se, essencialmente, em Imunologia, Genética e Biologia do Desenvolvimento. Nós, os “teóricos” – que habitavam os *dry labs* cheios de computadores - rondávamos uma dezena a trabalhar na ala dos Estudos Avançados de Oeiras.

Entre os “teóricos”, reinavam os físicos e os bioquímicos. Eu, com a minha formação em Estatística, sentia-me um estranho numa terra estranha, um verdadeiro *outlier*. Nem eu tinha ninguém por perto que percebesse, ou falasse, a minha linguagem estatística adquirida a custo no IST, nem eu entendia o que (me) diziam (genes RAG e HOX, intrões, exões e codões, citocinas e interleucinas, reacções de centros germinativos, os ratinhos NOD e B6, etc).

Fiquei impressionado pela constante actividade científica do Instituto, sob um clima de descontração e informalidade, mas não menos rigoroso e exigente, onde o Dr. Jorge Carneiro passou a ser apenas Jorge. Duas vezes por semana ocorriam, e continuam a ocorrer, os chamados “Seminários do IGC”: às terças-feiras, com investigadores da casa, e às sextas-feiras, com investigadores nacionais e internacionais provenientes de diversas áreas das Ciências da Vida. Havia, ainda, um seminário dos “teóricos”, e alguns “clubes de jornais” (de Imunologia e Genética), a completar este quadro de dinamismo institucional.

Seguindo as recomendações do Jorge, tentava estar sempre presente nessas reuniões, com o intuito de aprender rapidamente a linguagem das várias linhas de investigação do Instituto. Invariavelmente saía confuso desses eventos. Tantas moléculas, tantos modelos animais, tantas técnicas experimentais a saber!

Foi lento o processo de aprendizagem de uma nova linguagem, a da Biologia. Recebia frequentemente palavras de incentivo por parte dos meus colegas de grupo e da parte experimental. “É uma questão de tempo”, diziam-me. Mas, já diz o povo, “o que não nos mata, torna-nos mais forte”, e no meu caso, tornou-me mais conhecedor das dificuldades diárias de um “experimentalista”, não só em termos laboratoriais como estatísticos.

Ajudar para conquistar

Com o passar do tempo, senti a confiança a aumentar. Por um lado, começava a perceber o que os biólogos diziam, e queriam, quando me procuravam. Por outro lado, a minha linguagem parecia estar a tornar-se mais acessível a eles, evitando palavrões estatísticos, como homocedasticidade ou heterocedasticidade, que afugentava qualquer um que me perguntasse algo sobre a aplicação do teste T aos seus dados. “Bastou” saber quais os seus conhecimentos básicos em Estatística, para desenvolver uma estratégia de comunicação mais eficiente entre os dois mundos, o da Biologia e o da Estatística.

Assim, começaram a surgir colaborações com diversos investigadores, de dentro e fora do Instituto, e com elas, abriram-se novas avenidas na investigação em Estatística (aplicada). A primeira é sempre a primeira, e vale a pena referi-la, não só por interesses de narrativa histórica, mas também como um bom exemplo de uma colaboração bastante proveitosa.

Em 2003, o Dr. Pablo Pereira, um imunologista do Instituto Pasteur de Paris, pediu auxílio ao grupo de Imunologia Teórica na análise de duas tabelas de contingências sobre o estado de rearranjo de dois genes em células precursoras dos linfócitos T $\gamma\delta$. Por se tratarem de dados categorizados, um estatístico sentiria um apelo praticamente irresistível de recorrer aos tradicionais modelos (log-)lineares para esse tipo de dados. Porém, neste caso, a sua utilização conduziria a resultados que não dariam uma resposta cabal ao objectivo principal da análise: o de compreender o mecanismo de rearranjo desses genes. Optou-se, então, por uma abordagem estatística alternativa, tentando modelar o processo de rearranjo genético propriamente dito.

Sendo um novato em Imunologia, recebi uma atenção especial do Jorge na concepção dos modelos, que foram posteriormente implementados num formalismo de cadeias de Markov em tempo contínuo. A etapa de estimação revelou-se um pesadelo, devido à grande complexidade da função de verosimilhança, o que levou a usar-se o algoritmo *steepest descent*, que não requer o cálculo de derivadas, para maximizar essa função, e assim obter as respectivas estimativas para os parâmetros dos

modelos. No final da análise conseguiu-se ajustar um modelo com poucos parâmetros, que descrevia quase na perfeição os dados experimentais.

Comunicamos, então, os resultados ao nosso colaborador, e esperámos pacientemente que ele e o seu grupo de investigação conseguissem confirmar experimentalmente a validade do modelo ajustado, o que posteriormente o fizeram. Quando isso aconteceu, o trabalho realizado foi dividido em dois artigos, devido à sua extensão e complexidade: um de cariz experimental [1] e outro com a descrição pormenorizada da análise estatística efectuada [2].

Apesar de se ter conseguido publicar o trabalho resultante da colaboração, a maior lição a tirar deste episódio foi a de mostrar a necessidade de ter alguém que direcione a análise estatística no caminho certo. Eu, como estatístico, não tinha conhecimento imunológico suficiente para formular tais modelos. Teria ficado pela rama do problema, com o ajustamento dos modelos tradicionais para dados categorizados, o que não seria satisfatório. Neste caso, o Jorge Carneiro foi quem fez chegar esta colaboração a bom porto. Precisamos de mais investigadores com o seu perfil.

O estatístico como um paramédico científico

As colaborações foram surgindo a bom ritmo, mas não há bela sem senão. Estas ocorriam na maior parte das vezes em situações de “desespero”, quando os manuscritos submetidos para a publicação já tinham sido rejeitados pelos avaliadores, devido a alguma insuficiência estatística. Tornei-me num género de paramédico estatístico, aquele que respondia ao apelo de SOS, sempre que havia necessidade de reanimar os dados, que tinham entrado entretanto em coma, após os avaliadores terem posto o dedo na ferida.

Foi num cenário de “ser para ontem” que realizava muitas das análises estatísticas que me pediam. Tentava sempre manter uma certa calma, aquando da realização da análise estatística, pois esta precisa de amadurecer com o tempo, tal como o vinho. Queria evitar a todo o custo a minha contribuição pessoal para a actual tendência de crescimento de erros técnicos em artigos científicos [3]; espero tê-lo conseguido. Sentia uma certa renitência no ar, quando mencionava métodos “menos” conhecidos, mas potencialmente mais adequados para os fins inferenciais propostos. Notava que a expressão “apenas descrever os dados” causava amiúde algum desconforto a quem a ouvia, mas após o trabalho feito, havia uma concordância tácita entre todos em relação à abordagem estatística adoptada. Como um exemplo de sucesso foi a utilização da análise de correspondências na identificação de diferentes grupos de risco genético de raças ovinas portuguesas para a *scrapie*, a variante ovina da conhecida doença das vacas loucas [4].

Dificuldades em interpretar

Existem várias dificuldades que os estatísticos enfrentam quando têm um conjunto de dados a analisar. A primeira prende-se essencialmente com a qualidade dos dados: dimensões amostrais inadequadas, procedimentos experimentais incorrectos, existência de grandes quantidades de dados omissos, etc, etc. Como o estatístico continua a não ser consultado antes da realização de experiência, salvo em raras excepções, este não tem outra alternativa que não seja fazer um exame *post-mortem*, tal como Fisher se referia metaforicamente a situações deste tipo [5].

A seguir vem a dificuldade mais apetecível para um estatístico: a de encontrar a ferramenta estatística mais adequada para o problema em questão. Qual o melhor teste a aplicar aos dados? Que modelos se podem ajustar? Deve-se usar uma abordagem paramétrica, ou não-paramétrica? São apenas algumas questões que surgem ao longo de uma análise estatística propriamente dita. Por vezes, os métodos que se tem a disposição não são suficientes para atingir os objectivos propostos no início do estudo. E assim começa a investigação (em) estatística. Com um pouco de sorte, desenvolve-se um novo método de análise. Um novo artigo a (tentar) publicar?

Por fim, chega-se a fase de interpretação de resultados. E aqui surge mais uma dificuldade. “A estatística não prova nada! Por isso muito cuidadinho”, lembro-me do Prof. Daniel Paulino me dizer inúmeras vezes, durante a realização da minha tese de mestrado. Neste aspecto, há que ter sempre ciente que um resultado estatístico pode não ter qualquer tradução biológica.

Aqui vão dois exemplos práticos em que senti essa dificuldade interpretativa na pele.

Em 2004, comecei a cooperar com o Dr. Carlos Penha-Gonçalves, o líder do grupo de Doenças Genéticas do IGC. Ao saber da minha formação em Estatística, e da minha necessidade de ter um tema para a minha tese de mestrado, propôs-me um trabalho muito interessante sobre a modelação da interacção genética, ou *epistasis*, em fenótipos binários complexos. Ele e o seu grupo de investigação tinham-se deparado com dificuldades em interpretar o bom ajustamento de vários modelos lineares generalizados a um conjunto de dados referentes à interacção de duas zonas do genoma no controlo da susceptibilidade à malária cerebral em ratinhos [6]. De facto, o uso desses modelos em dados genéticos já tinha sido questionado, pela fraca interpretação genética que deles se pode retirar [7]. Para complicar este cenário, não parecia existir um consenso generalizado na literatura genética sobre a verdadeira natureza (estatística ou biológica) do termo *epistasis* [8].

Com este ponto de partida, tentou-se abordar o mesmo problema inferencial, mas com uma visão mais focada na Biologia. Sob a batuta do Daniel Paulino e do Carlos Penha-Gonçalves, que orientaram a minha tese, e dos comentários acutilantes, mas sempre pertinentes, do Jorge Carneiro e dos meus colegas de grupo (obrigado, Rui e Tiago), conseguiu-se desenvolver uma nova abordagem à modelação de interacção genética baseada no conceito de expressão alélica, que parece ter mais informação biológica do que os tradicionais modelos lineares generalizados [9-11].

Mais tarde, já em pleno doutoramento, debati-me com outro problema interpretativo, quando tentava localizar zonas do genoma que estivessem a influenciar as proporções de diferentes populações de timócitos em ratinhos [12]. Ao aplicar a análise de componentes principais aos dados, foi possível identificar três regiões candidatas. Apesar dessas regiões conterem genes descritos na literatura que parecem afectar efectivamente o número de timócitos, a natureza biológica das componentes principais encontradas foi desde logo questionada, atendendo às perigosas interpretações que esta e outras técnicas multivariadas proporcionaram, num passado recente, sobre matérias tão delicadas, como são as relacionadas com o carácter hereditário da inteligência humana [13]. Do meu ponto de vista, uma resposta cabal ao problema de interpretação das componentes principais para o número de timócitos só poderá ser dada com o desenvolvimento de uma abordagem estatística, que incorpore os mecanismos de acção dessas células durante o seu desenvolvimento no timo, o que até agora não foi feito.

A actualidade

Nos últimos quatro anos o IGC cresceu em número de investigadores, assim como em temáticas de investigação. Vários programas de doutoramento (Biologia Computacional, Neurociências, entre outros) foram implementados, trazendo mais gente ao Instituto. Os “teóricos” passaram a ser mais de trinta, distribuídos por diversos grupos de investigação. Em relação ao grupo de Imunologia Teórica, transformou-se no actual grupo de Biologia Quantitativa do Organismo, sob a mesma liderança de sempre (a do Jorge Carneiro).

Este crescimento trouxe consigo novos desafios científicos com fortes necessidades estatísticas. Assim, parece existir, hoje em dia, uma janela de oportunidade para a afirmação da Estatística dentro do Instituto. Sem dar conta, dou por mim a pensar na importância fulcral que a Estação Agronómica de Rothamsted no Reino Unido teve no desenvolvimento da Estatística como ciência, na primeira metade do século XX [14]. E ao escrever isto, apercebo-me que o IGC habita paredes-meias com a Estação Agronómica Nacional...

A grande diversidade de pessoas do Instituto fez-me também confrontar com trabalhos aplicados de grande complexidade estatística, executados por investigadores sem uma formação básica na área. Não querendo parecer injusto para com muitos desses investigadores (nomeadamente, Isabel Gordo, José Pereira Leal e Gabriela Gomes), vou apenas destacar o trabalho do grupo de Genética da Conservação e das Populações, encabeçado por Lounès Chikhi, pelo meu mero gosto pessoal pela sua temática de investigação.

Esse grupo tem como objectivo a reconstrução do historial demográfico de populações através da análise de dados genéticos. Diversos métodos baseados na teoria da coalescência têm sido utilizados, com o intuito de estimar parâmetros de diferentes modelos demográficos. A título ilustrativo, na área da conservação, interessa detectar, quantificar e datar eventos de redução demográfica drástica (*bottlenecks*), ou de separação de populações. As inferências a realizar nesses modelos são complexas e computacionalmente exigentes, sendo as melhores respostas dadas por métodos baseados na

utilização massiva de simulação estocástica, tal como os MCMC e ABC (*approximate bayesian computation*) [15]. A este respeito, o grupo tem desenvolvido novas metodologias de análise, assim como perceber as limitações das actuais ferramentas inferenciais [16,17].

Olhando o horizonte

Em Agosto de 1998 o cientista norte-americano Craig Venter, que liderou a iniciativa privada do Projecto do Genoma Humano, afirmava na revista *Business Week* que o século XXI seria dedicado à Biologia. E de facto, é o que parece estar a acontecer.

Actualmente estamos a assistir a uma explosão de projectos em ciências biológicas, ou biomédicas, impulsionados pelo advento desse grande empreendimento relacionado com a sequenciação do genoma humano [18,19] e da criação do HapMap, um mapa de haplótipos em diferentes populações humanas [20,21]. Estes dois grandes projectos criaram condições muito favoráveis para, por exemplo, localizar, com mais exactidão, zonas do genoma “responsáveis” por muitas doenças humanas comuns através da utilização de estudos de associação [22]. Nova metodologia estatística é, no entanto, necessária para fazer um uso eficiente de toda a informação disponível [23], o que já começa a ser feito [24,25].

Os últimos dez anos têm também mostrado a Biologia como uma ciência em constante desenvolvimento tecnológico. Novas técnicas experimentais têm surgido quase a um ritmo diário, o que conduziu a novos desafios em termos de análise estatística. A este respeito, as tecnologias de grande débito de dados, tal como os conhecidos *microarrays* ou *genechips*, com a medição simultânea da actividade de milhares de genes num conjunto reduzido de indivíduos, estão na ordem do dia. Questões sobre o correcto delineamento experimental a aplicar a este tipo de experiências, ou como comparar eficazmente vários grupos experimentais medidos por estas tecnologias, são apenas dois tópicos que os estatísticos terão certamente uma palavra a dizer, pelo que se pode julgar das discussões que se encontram na literatura [26]. Contudo, a integração estatística da informação proveniente de diferentes procedimentos experimentais é ainda escassa [27].

Posto isto, o panorama actual da Biologia parece reservar aos estatísticos um futuro auspicioso. E porque não dizê-lo, para os matemáticos também [28]. Porém, os estatísticos (e os matemáticos) terão de estar dispostos a criarem laços fortes com os seus colegas biólogos.

Apesar da minha ainda curta experiência profissional, penso que a chave do sucesso futuro reside, essencialmente, em duas palavras: colaboração e integração.

Por um lado, uma forte colaboração entre estatísticos e biólogos é necessária para que haja a possibilidade de encontrar novos, e interessantes, problemas estatísticos. Para isso, sugiro que os estatísticos (académicos) se dêem a conhecer, ora assistindo a seminários, ou conferências, em alguma área da Biologia que lhes interesse, ora convidando, com alguma regularidade, biólogos ou investigadores a trabalhar no interface das duas disciplinas, para darem palestras nos seus departamentos.

Por outro lado, os estatísticos devem ser capazes de se integrar em equipas multidisciplinares, onde várias áreas do conhecimento estão representadas. Sendo por vezes difícil, a conjugação de diferentes pontos de vista, mesmo que pareçam distantes da Estatística, é uma etapa que considero actualmente crucial numa análise estatística (obrigado Jorge, por me teres feito ver isto).

Acerca deste assunto, a minha visão, possivelmente enviesada pela formação, ou deformação, académica que recebi, tende a mostrar-me os estatísticos a tentarem encontrar modelos simples, mas de cariz puramente matemático, que se ajustem bem aos dados, enquanto que os físicos e os biólogos a pensarem mais em termos de mecanismos hipotéticos de acção, que possam explicar os fenómenos observados. A conjugação destas duas visões poderá, então, produzir metodologias estatísticas inovadoras, que atinjam a tão desejada interpretação biológica, mas que ao mesmo tempo, obedeçam aos princípios teóricos que regem uma boa análise estatística.

Seguindo de perto esta filosofia, ando em mãos com um problema de forte índole estatística: a estimação de diversidade de linfócitos T [29]. Ao analisar dados disponíveis na literatura, dois modelos parecem ajustar-se igualmente bem aos dados observados: o Lognormal-Poisson e Yule. Tendo este resultado, chega-se à pergunta fulcral da análise: qual o melhor modelo para os dados?

Pelo princípio da parcimónia, um estatístico seria tentado a preferir o modelo Yule, pois este tem menos parâmetros do que o seu rival Lognormal-Poisson. Contudo, ao trabalhar com um modelo que descreve a dinâmica de diferentes populações de linfócitos T no organismo [30], as respectivas simulações parecem apontar para a melhor adequação do modelo Lognormal-Poisson (resultados a aparecer na minha tese de doutoramento).

Para mim, este pequeno exemplo demonstra bem o benefício que a Estatística tem em se aliar a outras abordagens para o mesmo problema inferencial.

Depois de tantas divagações resta-me concluir.

Neste momento trabalho num limbo entre a Estatística e a Biologia, tal como o Tiago Marques. Com o tempo, vou-me tornar cada vez mais um sofrível estatístico, e temo que nunca, por mais que me esforce, serei um biólogo, embora haja quem diga o contrário. Mas sem dúvida, sinto-me cada vez mais capacitado em fazer a ligação entre as duas áreas. A investigação estatística apenas tem a ganhar da cooperação com a Biologia, ou com vários outros domínios científicos (e.g., Medicina, Física ou Ciências Sociais). É necessário, porém, que os estatísticos estejam dispostos a “sair da casca”, e dar o salto para o outro lado, e que os biólogos (ou outros cientistas) tenham o discernimento, e a vontade, de procurar a colaboração numa fase inicial do seu trabalho. Mãos à obra!

Agradecimentos

Ao Fernando Rosado por me ter convidado a escrever este texto. Ao Tiago Marques pela troca de ideias (“great minds think alike...”, LOL). Ao IGC e a sua “gente”, por ser um excelente local para trabalhar e fazer ciência. Ao Jorge Carneiro (IGC), Carlos Penha-Gonçalves (IGC) e Daniel Paulino (IST) pelas suas constantes palavras de apoio e incentivo ao longo destes últimos 8 anos. Ao Lounès Chikhi (IGC) e Vítor Sousa (IGC) pela sua disponibilidade em discutir os seus trabalhos comigo.

Referências

- [1] Boucontet, L., Sepúlveda, N., Carneiro, J. and Pereira, P. (2005). Mechanisms controlling termination of V-J recombination at the TCR γ locus: implications for allelic and isotypic exclusion of TCR γ chains. *J. Immunol.* **174**:3912–3919.
- [2] Sepúlveda, N., Boucontet, L., Pereira, P. and Carneiro, J. (2005). Stochastic modeling of T cell receptor γ gene rearrangement. *J. Theor. Biol.* **234**:153–165.
- [3] Hawkins, B. A. (1999). More haste, less science? *Nature* **400**:498.
- [4] Orge, L., Galo, A., Sepúlveda, N., Simas, J. P. and Pires, M. (2003). Scrapie genetic susceptibility in Portuguese sheep breeds. *Vet. Rec.* **153**:508.
- [5] Fisher, R. A. (1962). The place of the design of experiments in the logic of scientific inference. *Colloq. Intern. Cent. Nat. Recherche Scientifique*, **110**:13-19.
- [6] Bagot, S., Campino, S., Penha-Gonçalves, C., Pied, S., Cazenave, P. A. and Holmberg, D. (2002). Identification of two cerebral malaria resistance loci using an inbred wild-derived mouse strain. *Proc. Natl. Acad. Sci. USA* **99**:9919-9923.
- [7] Cordell, H., Todd, J., Hill, N., Lord, C., Lyons, P., Peterson, L., Wicker, L. and Clayton, D. (2001). Statistical Modeling of Interlocus Interactions in a Complex Disease: Rejection of the Multiplicative Model of Epistasis in Type I Diabetes. *Genetics* **158**:357-367.
- [8] Moore, J. H. and Williams, S. M. (2005). Traversing the conceptual divide between biological and statistical epistasis: systems biology and a more modern synthesis. *Bioessays* **27**:637–646.
- [9] Sepúlveda, N. (2004). Modelos estatísticos para a ação conjunta de dois loci em fenótipos binários complexos. Tese de Mestrado, Departamento de Matemática, Instituto Superior Técnico.
- [10] Sepúlveda, N., Paulino, C. D., Carneiro, J. and Penha-Gonçalves, C. (2007). Allelic penetrance approach as a tool to model two-locus interaction in complex binary traits. *Heredity* **99**:173-184.
- [11] Sepúlveda, N., Paulino, C. D. and Penha-Gonçalves, C. (2009). Bayesian analysis of allelic penetrance models for complex binary traits. *Comp. Stat. Data Anal.* **53**:1271-1283.
- [12] Sepúlveda, N. e Carneiro, J. (2007). Que loci controlam as proporções das diversas populações de timócitos? Em *Estatística: Ciência Interdisciplinar* (Ferrão, M. E., Nunes, C. e Braumann, C. A., eds), 729-740, Edições SPE, Lisboa.

- [13] Gould, S. J. (1996). *The Mismeasure of Man*. 2ª Edição, W. W. Norton, Nova Iorque.
- [14] Fisher, R. A. (1933). The contributions of Rothamsted to the development of the science of Statistics. Annual Report of the Rothamsted Experimental Station, 43-50
- [15] Beaumont, M. A. and Rannala, B. (2004). The Bayesian revolution in Genetics. *Nat. Rev. Genet.* **5**:251-261.
- [16] Chikhi, L., Bruford, M. W. and Beaumont, M. A. (2001) Estimation of admixture proportions: a likelihood-based approach using Markov chain Monte Carlo. *Genetics* **158**:1347-1362.
- [17] Sousa, V., Fritz, M., Beaumont, M. and Chikhi, L. (2009). Approximate Bayesian Computation without summary statistics: the case of Admixture. *Genetics* (no prelo).
- [18] International Human Genome Sequencing Consortium (2001). Initial sequencing and analysis of the human genome. *Nature* **409**:860-921.
- [19] International Human Genome Sequencing Consortium (2004). Finishing the euchromatic sequence of the human genome. *Nature* **431**:931-45.
- [20] International HapMap Consortium. (2005). A haplotype map of the human genome. *Nature* **437**:1299-1320.
- [21] International HapMap Consortium. (2007). A second generation human haplotype map of over 3.1 million SNPs. *Nature* **449**:851-861.
- [22] Manolio, T. A., Brooks, L. D. and Collins, F. S. (2008). A HapMap harvest of insight into the genetics of common disease. *J. Clin. Invest.* **118**:1590-1605.
- [23] Zhang, W., Ratain, M. J. and Dolan, M. E. (2008). The HapMap resource is providing new insights into ourselves and its application to pharmacogenomics. *Bioinform. Biol. Insights* **2**:15-23.
- [24] Liu T, Johnson, J. A., Casella, G. and Wu, R. (2004). Sequencing complex diseases with HapMap. *Genetics* **168**:503-11.
- [25] Wu, S., Yang, J., Wang, C. and Wu, R. (2007). A General Quantitative Genetic Model for Haplotyping a Complex Trait in Humans. *Curr. Genomics* **8**:343-350.
- [26] Akey, J. M., Biswas, S., Leek, J. T. and Storey, J. D. (2007). On the design and analysis of gene expression studies in human populations. *Nat. Genet.* **39**:807-808.
- [27] Ge, H., Walhout, A. J. M. and Vidal, M. (2003). Integrating 'omic' information: a bridge between genomics and systems biology. *Trends Genet.* **19**:551-560.
- [28] Cohen, J. E. (2004). Mathematics is Biology's next microscope, only better; Biology is Mathematics' next physics, only better. *PLoS Biol.* **2**:e439.
- [29] Sepúlveda, N., Paulino, C. D. e Carneiro, J. (2008). Diversidade de linfócitos T reguladores e efetores. Em *Estatística: da Teoria à Prática* (Hill, M. M., Ferreira, M. A., Dias, J. G., Salgueiro, M. F., Carvalho, H., Vicente, P. e Braumann, C. A., eds), 513-524, Edições SPE, Lisboa.
- [30] Carneiro, J., Léon, K, Caramalho, I., van den Dool, C., Gardner, R., Oliveira, V., Bergman, M. L., Sepúlveda, N., Paixão, T., Faro, J. and Demengeot, J. (2007). When three is not a crowd: a Crossregulation model of the dynamics and repertoire selection of regulatory CD4+ T cells. *Immunol. Rev.* **216**:48-68.



Estatística na Análise do Território

Susana Rosado Ganhão*, *srosado@fa.utl.pt*

Cristina Delgado Henriques**, *cdh@fa.utl.pt*

Faculdade de Arquitectura da UTL

** - Centro de Estatística e Aplicações da FCUL*

*** - e-GEO Centro de Estudos de Geografia e Planeamento Regional da FCSH/UNL*

1. Enquadramento

Os técnicos envolvidos no processo de planeamento e ordenamento do território deverão possuir uma visão integrada do meio físico e dos sistemas humanos que nele se localizam. Consequentemente, necessitam de um conhecimento profundo das características fisiográficas do território, dos aspectos socio-económicos das populações, bem como das consequências resultantes das várias opções possíveis no que diz respeito a transformações no uso do solo.

O estabelecimento de indicadores para monitorizar e avaliar acções de planeamento e ordenamento do território é actualmente um requisito para o exercício pleno da cidadania. Do ponto de vista dos decisores o estabelecimento de indicadores permitirá decisões informadas, escolhas adequadas e intervenções racionalizadas; do lado dos cidadãos, exercer um controlo documentado relativo às acções sobre o território.

O acesso fácil a informação variada e actualizada relativa a diversos sectores e disciplinas, assim como a disponibilidade de meios de análise dessa informação, parecem, portanto, ser cada vez mais imprescindíveis num processo de planeamento e ordenamento do território sustentável.

Os Sistemas de Informação Geográfica (SIG) são instrumentos que ao integrarem numa única base de dados informação de natureza diversa, nomeadamente descritiva e gráfica, e estando dotados de ferramentas de análise espacial que permitem a simulação de cenários alternativos, vieram introduzir um novo horizonte aos técnicos e responsáveis pelo ordenamento do território. “... *GIS is beyond mapping and has been increasingly used in various spatial analysis tasks as the GIS software becomes more capable and also friendlier to use for these tasks.*” (Wang, 2006).

A Detecção Remota (DR) oferece também um conjunto de vantagens para o ordenamento do território ao disponibilizar um conjunto de dados integráveis em bases de dados georreferenciadas e ao proporcionar uma visão do território que possibilita a identificação de padrões e relações que não são perceptíveis do solo.

A compreensão da distribuição espacial de dados geográficos passa, para além da percepção visual da distribuição dos problemas proporcionada pelos SIG e pela DR, pela quantificação, pelo estabelecimento de relações e padrões existentes, pela generalização e pela observação de processos através do tempo para formular previsões.

Um processo de decisão que envolva dados espaciais pode ter um ou mais dos seguintes objectivos (Fischer et al, 1996):

- procura de padrões;
- interpretação de fenómenos naturais ou sociais;
- gestão de informação;
- predição da localização e intensidade do fenómeno em estudo.

A análise de dados espaciais é uma das funcionalidades dos SIG. Estes sistemas possuem uma vasta gama de ferramentas analíticas desde as mais básicas (medição, classificação, sobreposição, vizinhança, conectividade) às mais avançadas que envolvem o estudo estatístico e matemático de dados baseado em modelos teóricos (amostragem, estatística descritiva, análise multivariada, análise de séries temporais, análise geostatística, análise bayesiana, modelação hierárquica, optimização e simulação).

Goodchild et al (1992) invocam um conjunto de pontos que evidenciam a importância da componente espacial dos dados:

1. o espaço fornece um indexamento conveniente;
2. a perspectiva espacial permite um acesso fácil à informação;
3. permite associação de diferentes eventos num SIG, através de *layers*;
4. permite que a distribuição espacial dos objectos e fenómenos sejam usados para representar interacções espaciais.

Nos pontos seguintes apresentam-se duas aplicações concretas da estatística à análise territorial e que constituem um contributo para o planeamento e ordenamento do território.

2. Identificação de *clusters* urbanos

Um breve olhar sobre uma periferia de uma cidade africana pode levar-nos à percepção da existência de uma aparente homogeneidade entre os diversos bairros que a compõem (Figura 1). Porém, uma observação mais pormenorizada, auxiliada por trabalho de campo, trouxe algumas questões: terão os bairros do subúrbio, contemporâneos do período colonial, os mesmos padrões habitacionais que os bairros que surgiram durante e depois da década de 70 do século XX? Será idêntica a forma e a densidade de ocupação residencial do solo nos bairros mais próximos do centro e dos bairros mais afastados? Ou seja, será que existe diversidade dentro de uma aparente homogeneidade?

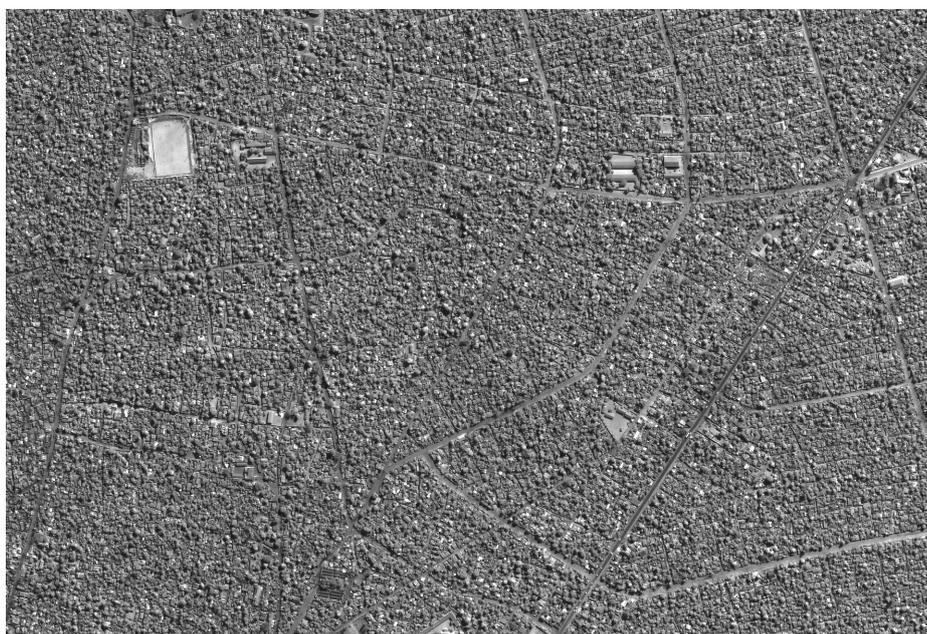


Figura 1: Extracto de uma imagem do satélite IKONOS relativa à periferia da cidade de Maputo

A partir de um conjunto de variáveis que descrevem as características sócio-económicas e habitacionais da população da periferia de uma cidade africana (Maputo) utilizou-se (em Henriques, 2008) a Análise em Componentes Principais (ACP) como técnica exploratória que possibilitou, numa primeira fase, estabelecer ligações entre as variáveis envolvidas integrando a sua influência mútua, e evidenciar semelhanças/parecenças e dissemelhanças/diferenças entre bairros traduzidas numa descrição tipológica. Através de um método hierárquico de Classificação Automática (CAH), a que estão subjacentes critérios de similitude e de dissimilitude das variáveis relativas à habitação, classificaram-se os bairros em grupos (Figura 2).

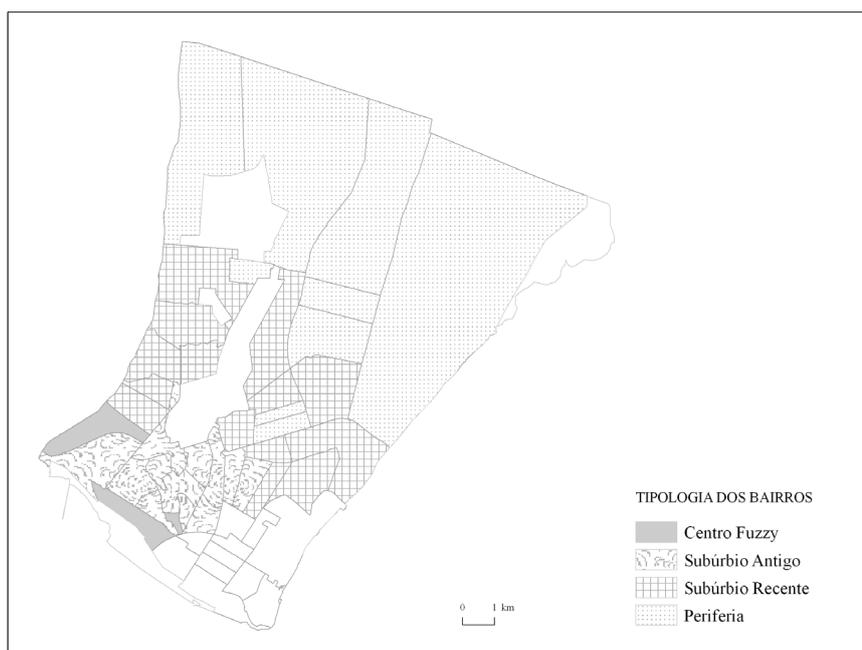


Figura 2: Tipologia de bairros da cidade de Maputo resultante da utilização da ACP e da CAH (adaptado de Henriques, 2008)

Com recurso a estas ferramentas estatísticas foi possível evidenciar a existência notória de grupos (*clusters*) que permitiram definir uma tipologia dos bairros da cidade em: subúrbio antigo, subúrbio recente, periferia e zona indefinida (*fuzzy*), disponibilizando aos actores do planeamento urbano informação para uma intervenção diferenciada nos bairros da cidade de Maputo.

3. Geoestatística

A geoestatística, segundo Soares (2006), é a denominação de um conjunto de métodos, técnicas e instrumentos estatísticos que caracterizam os fenómenos espaciais naturais.

Planear e gerir um recurso natural com vista à sua preservação ou recuperação, como por exemplo uma floresta ou um aquífero, implica geralmente um primeiro passo fundamental que consiste em quantificar e caracterizar os fenómenos espaciais implícitos naqueles recursos.

A carência de informação e a variabilidade do fenómeno constituem as razões da nossa ignorância em relação a este, i.e., estão na base da incerteza associada ao seu conhecimento. É, precisamente, este o objecto da geoestatística: a caracterização dos fenómenos espaciais naturais e a quantificação da incerteza do seu conhecimento.

A sua metodologia consiste, resumidamente, na quantificação estatística da continuidade espacial da grandeza em estudo, em modelos de interpretação espacial tendo por base a sua variabilidade estrutural e em modelos de simulação estocástica que quantificam a incerteza ligada ao fenómeno espacial.

Alguns exemplos ilustrativos de estudo geoestatístico são apresentados em Soares (2006): caracterização espacial de lodos de um rio, tendo por objectivo a sua remediação, Pereira et al (1999); definição do padrão espacial de aves migratórias (rola) em Portugal Continental, Santos (1998); estudo da dispersão espacial da população de duas espécies protegidas de limículas para avaliação do impacte da construção da Ponte Vasco da Gama, Pereira et al (1998); distribuição espacial da idade e da qualidade no inventário do eucalipto em Portugal Continental, Cresp et al (1998); análise dos custos de impacte da poluição atmosférica na península de Setúbal, Nunes (1998); dispersão espacial de um índice de vitalidade do sobreiro no Baixo Alentejo, Mateus et al (1998); caracterização espacial da salinidade troposférica, Figueira et al (1998).

4. Considerações finais

A análise do território sob as suas diferentes dimensões, físicas, sociais, económicas, culturais... envolve o conhecimento de um grande número de variáveis importantes em cada uma dessas dimensões. Estabelecer relações e encontrar explicações são questões fundamentais quando se pretende intervir de forma sustentada no território.

O desenvolvimento tecnológico tem apoiado muitas destas intervenções disponibilizando ferramentas de aquisição, tratamento e análise de grandes volumes de dados espaciais. A integração da estatística nestas ferramentas proporciona um conhecimento mais profundo e articulado da realidade onde se pretende intervir através do estabelecimento de modelos inferenciais.

Uma das técnicas estatísticas que tem vindo a ser muito usada para análise de dados espaciais é a modelação hierárquica, incluindo os métodos bayesianos hierárquicos, que dão dividendos substanciais na estatística espacial e espaço-temporal.

A evolução tecnológica na implementação (computacional) de métodos como os MCMC (Markov Chain Monte Carlo) permitiu que fossem possíveis análises bayesianas completas de sofisticados modelos com múltiplos níveis para dados complexamente georeferenciados (Banerjee et al, 2004). Este tipo de abordagem também permite inferência não Gaussiana de dados espaciais, dados espaciais multivariados, dados espaço-temporais e, inclusivé, soluções para problemas tais como desalinhamentos geográficos e temporais de *layers* de dados espaciais.

A análise do território é, assim, uma fonte promissora para o desenvolvimento de técnicas estatísticas aliadas aos SIG, que permitam suportar a tomada de decisões relacionadas com a gestão da informação espacial georreferenciada.

Referências

- BANERJEE, Sudipto, Carlin, Bradley P., Gelfand, Alan E. (2004) *Hierarchical Modeling and Analysis for Spatial Data*. Chapman & Hall/CRC, New-York.
- CRESPO, R., Fonseca, L., Rocharte, L.(1998) Distribuição Espacial da Idade e da Qualidade no Inventário do Eucalipto em Portugal Continental, *Actas do 5º Seminário Interno do CMRP/IST*, Castelo de Vide, Portugal.
- FISCHER, M., Scholten, H.J., Unwin, D.(1996) *Spatial Analytical Perspectives on GIS*. Taylor & Francis. 256 p.
- GOODCHILD, M., Haining, R., Wise, S. (1992) Integrating GIS and spatial data analysis: problems and possibilities. *International Journal of Geographic Information Systems*, 6(5), 407-23. London: Taylor & Francis
- HENRIQUES, C. (2008) *Maputo. Cinco Décadas de Mudança territorial. O uso do solo observado por tecnologias de informação geográfica*. Instituto Português de Apoio ao Desenvolvimento. p.238
- MATEUS, F., Ravasco, A. (1998) Caracterização da Dispersão Espacial de um Índice de Vitalidade do Sobreiro no Baixo Alentejo. *Actas do 5º Seminário Interno do CMRP/IST*, Castelo de Vide, Portugal.
- NUNES, C.(1998) Caracterização Espacial de um Poluente Atmosférico Resultante de Emissões Multi-fonte, Tese de Mestrado, IST.
- PEREIRA, M.J., Almeida, J., Rodrigues, A., Moura, D.(1999) Mapping Spatial Distributions of Waders on South Tagus Riverside, *geoENV II – Geostatistics for Environmental Applications*. Gomes,-Hernandez et al., Eds., Kluwer Academic Pub., 225-236.
- SANTOS, E.(1998) Caracterização Geostatística da Distribuição Espacial da Rola em Portugal Continental, *Actas do 5º Seminário Interno do CMRP/IST*, Castelo de Vide, Portugal.
- SOARES, A. (2006) *Geoestatística para as Ciências da Terra e do Ambiente*. 2ª Edição. IST Press. 214 p.
- WANG, F. (2006) *Quantitative Methods and Applications in GIS*. Taylor & Francis. 265p.
- WONG, D.; Lee, J. (2005) *Statistical Analysis of Geographic Information*. John Wiley & Sons. 441p.



Avaliação contínua e testes com itens de escolha múltipla - algumas considerações

Patrícia Costa*, *pcosta@ubi.pt*
Pedro Nuno Oliveira**, *pno@dps.uminho.pt*
Maria Eugénia Ferrão*, *meferrao@ubi.pt*

* - *Universidade da Beira Interior*

** - *Universidade do Minho*

No início da década de 70 o número de alunos nas Instituições de Ensino Superior (IES) era aproximadamente de 60 mil e no ano lectivo 2007/8 cerca de 377 mil, não incluindo os alunos matriculados em cursos de pós-graduação. A expansão verificada aumentou a heterogeneidade da população discente, quer quanto ao perfil e trajectória escolar, quer quanto à origem socioeconómica, entre outros. Tal heterogeneidade não é compatível com a uniformização das metodologias de ensino e aprendizagem tradicionalmente usadas. A massificação do acesso assentou fundamentalmente na criação de lugares discentes, com preocupação insuficiente pela qualidade da formação, tendo como resultado a situação actual de performance do sistema – elevada taxa de abandono e de retenção, principalmente nos primeiros anos. Alguns dos factores de insucesso escolar no Ensino Superior apontados na literatura referem os conteúdos, o nível de compreensão alcançado, a descontinuidade em relação às experiências educativas anteriores (no domínio pedagógico, psicológico, didáctico e institucional) (Tavares, Santiago e Lencastre, 1998), a relação entre a atitude do professor e a motivação do aluno, o desempenho e trajectória escolar do aluno no Ensino Secundário (Cosme, 2004; Miranda e Marques, 1996), o desajuste entre níveis de ensino, a falta de bases, a insuficiência de recursos didácticos, a má organização de horários e calendários de testes, a falta de flexibilidade curricular, a falta de articulação vertical e horizontal, a deficiente integração entre teoria e prática e demasiado tempo dispendido em avaliações (Alarcão, 2000).

É neste contexto que novos desafios se colocam às IES portuguesas. A implementação das reformas decorrentes da Declaração de Bolonha, e preconizadas no Decreto-Lei nº42/2005, impõe novas necessidades em termos da prática docente, do material didáctico utilizado e das condições de infra-estruturas e de equipamento. Tal como o diploma refere, nesta nova concepção de ensino-aprendizagem, “o estudante desempenha o papel central, quer na organização das unidades curriculares, cujas horas de contacto assumirão a diversidade de formas e metodologias de ensino mais adequadas, quer na avaliação e acreditação, as quais considerarão a globalidade do trabalho de formação do aluno, incluindo as horas de contacto, as horas de projecto, as horas de trabalho de campo, o estudo individual e as actividades relacionadas com a avaliação, (...)”.

Nestes termos, a aferição das aprendizagens ganha relevância como parte integrante do processo ensino-aprendizagem (Petocz e Reid, 2007). No âmbito da Declaração de Bolonha, algumas IES alteraram e reorganizaram os seus currícula evidenciando que é uma excelente oportunidade para rever as metodologias de ensino e de aferição das aprendizagens (Teixeira, Silva e Flores, 2007).

Soeiro e Cabral (2004) propuseram o conceito de avaliação distribuída que é baseada numa mistura de vários métodos de avaliação, por exemplo, trabalho em laboratórios, mini-testes aplicados durante as aulas, trabalhos de casa ou projectos realizados individualmente ou em grupo. Existem outros métodos de aferição das aprendizagens como ensaios, perguntas de resposta aberta e testes com itens de escolha múltipla (TEM).

As linhas de orientação da European Association for Quality Assurance in Higher Education (ENQA, 2005, p.17) para a aferição das aprendizagens incluem os procedimentos desejáveis a serem implementados pelas instituições para ultrapassar a abordagem amadora e pouco precisa que é habitual. Os procedimentos vêm ao encontro das recomendações que decorrem de algumas investigações realizadas. A eficácia da aferição das aprendizagens dos alunos tem que reflectir os conteúdos do programa (Jolliffe, 1997; Holmes 2002) e deve ser válida e fiável (Wakeford, 2003).

A validade é o conceito que traduz a ideia do quão bem o teste está a medir o atributo ou constructo que é suposto medir. Nesse sentido, a validade refere-se tanto à evidência como à teoria com as quais as classificações no teste são passíveis de interpretação face ao propósito ou objectivo desse teste. Deste modo, a validade está associada à interpretação das classificações e ao seu uso e não ao próprio teste (AERA, APA, NCME 1999).

A fiabilidade de um procedimento de medição, neste caso o teste, é a capacidade que o procedimento tem para retornar medidas consistentes do atributo ou constructo quando as medições são repetidamente obtidas (réplicas), sob condições idênticas (Dunn, 1989). Constructo, do termo anglo-saxónico *construct*, é um meio intelectual para “construir um acontecimento” (Cronbach, 1951) e segundo Nunnally e Durham (1975, p.305) é uma denominação, que não tem existência real no mundo observável, sendo um instrumento heurístico para explorar o que se observa.

Os TEM têm sido utilizados cada vez mais como instrumento para a avaliação formativa no sentido de implementar a avaliação contínua, servindo assim para balizar as aprendizagens dos alunos. A utilização deste tipo de testes permite fornecer aos alunos informação sobre o seu progresso ao longo da unidade curricular, constituindo um factor motivador das aprendizagens e, nesse sentido, influencia o próprio processo de aprendizagem. As principais vantagens da utilização dos TEM são: maior cobertura dos programas das unidades curriculares; maior facilidade de aplicação, apesar de requerem maior tempo de preparação; maior uniformidade nos critérios de correcção com consequente redução de erro de classificação dos itens; maior rapidez na publicação dos resultados e a construção de um banco de itens, a partir do qual, futuros testes podem ser elaborados.

As limitações na utilização deste tipo de testes são: a dificuldade de adaptação a questões de resposta fechada (escolha múltipla, verdadeiro/falso, ou similares) em algumas disciplinas; as vantagens comparativas das perguntas de resposta aberta relativamente a perguntas de resposta fechada; a dificuldade em construir bons itens que possam expor e desenvolver ideias mais estruturadas.

Importa, por último, referir que os TEM exigem um esforço de controlo de qualidade das questões, controlo que tem que ser assegurado, entre outros, por métodos estatísticos. Obviamente, subsiste a necessidade de manter o banco de itens seguro.

Também Hand (2004, p.135) entende que a fiabilidade do procedimento de medição é um indicador do erro de medição, sendo o coeficiente de fiabilidade usado para a sua quantificação. Cronbach (1951, p.306) sugere a existência de diversos coeficientes para quantificar a fiabilidade e propõe o coeficiente alfa, coeficiente de consistência interna, para a obtenção duma estimativa: “ α is therefore an estimate of the correlation expected between two tests drawn at random from a pool of items like the items in the test”. Este coeficiente é de uso mais generalizado no âmbito da Teoria Clássica dos Testes TCT (Lord and Novick, 1968). Outras estatísticas muito utilizadas, para a caracterização das propriedades de um instrumento de medição são os índices de discriminação e dificuldade dos itens (o índice de discriminação mede a capacidade do item diferenciar os alunos com elevado desempenho daqueles que têm menor desempenho no teste; o índice de dificuldade é dado pela proporção de acerto no item, sendo, assim, uma escala invertida), correlação ponto bisserial, correlação item-total.

Quando o propósito é a comparação dos resultados da aferição das aprendizagens ao longo de anos sucessivos, a TCT impõe a utilização de testes paralelos ou do mesmo teste. Estes factos constituem uma grande limitação, não só pela dificuldade na concepção de testes paralelos, mas também pela dificuldade de garantir a confidencialidade dos diversos testes produzidos.

Esta é uma das limitações que a classe de Modelos de Resposta ao Item - MRI (Birnbaum, 1968; Hambleton, Swaminathan and Rogers, 1991) ajuda a ultrapassar. Os MRI permitem constituir bancos de itens (calibrados), a criação de testes a partir do banco de itens e a construção de escalas para aferir aprendizagens. Assim, os itens/testes aplicados reunirão condições para a modelação dos dados com vista à obtenção de uma métrica comum que assegura a comparação dos resultados atingidos pelos alunos ao longo do tempo.

Steif e Dantzler (2005) quantificaram a compreensão de conceitos dos estudantes em Estática, utilizando TEM. O trabalho realizado teve como propósito identificar os conceitos fundamentais e os erros típicos dos estudantes nesta disciplina. Foi efectuada a análise estatística dos resultados dos alunos em cinco universidades e o instrumento mostrou-se válido e fiável (coeficiente de alfa de Cronbach de 0.89) para medir a aprendizagem conceptual de Estática. Num trabalho subsequente, Steif e Hansen (2006) realizaram comparações entre as classificações dos alunos em diferentes cursos. As evidências confirmaram a validade e fiabilidade do teste aplicado (coeficiente de alfa de Cronbach de 0.82).

No passado recente, realizaram-se em Portugal dois projectos de investigação cujo objectivo geral foi a melhoria dos métodos de aferição das aprendizagens na perspectiva do trinómio ensino - aprendizagem – avaliação. A finalidade última é melhorar o sucesso escolar, particularmente em disciplinas que tradicionalmente apresentam taxa de insucesso mais elevada, como por exemplo, Estatística em cursos para não-matemáticos.

Na Universidade da Beira Interior decorreu, em 2007 e 2008, o projecto de investigação intitulado Estatística e Matemática para o Desenvolvimento de Competências Profissionais (POCI/C/04.01.02/0155/0001/2006). O projecto visou o desenvolvimento da matriz de objectivos de aprendizagem, bem como o banco de itens associado às unidades curriculares de Estatística e Matemática envolvidas nos cursos de Química Industrial, Engenharia Química, Optometria, Arquitectura, Ciência Política e Relações Internacionais e Psicologia.

Em relação à unidade curricular de Estatística Aplicada à Psicologia II, a análise da qualidade dos testes aplicados baseou-se na abordagem TCT. Os resultados revelaram-se interessantes sob do ponto de vista da adopção de novos métodos de avaliação formativa, concretamente na utilização de TEM. Em cada ocasião de aferição, foram aplicados dois tipos de teste – teste computadorizado com itens de escolha múltipla e teste com perguntas de resposta aberta. Um intervalo de 15 minutos mediou a aplicação de um e de outro. Um dos resultados relevantes foi a forte correlação (0.85) entre as classificações dos estudantes em cada um dos dois tipos de teste. Esta correlação permite evidenciar que se existissem problemas de concepção no teste de escolha múltipla, o mesmo se verificaria no teste de resposta aberta (tradicionalmente usado). Tanto os resultados obtidos como a opinião dos estudantes sobre a adopção de testes computadorizados com itens de escolha múltipla são encorajadores do alargamento desta experiência a outras unidades curriculares/cursos no futuro [Ferrão (2008)].

Na Universidade do Minho está a decorrer outro projecto que visa a utilização, de forma complementar, da TCT e dos MRI para garantir a qualidade na aferição das aprendizagens na unidade curricular de Estatística do curso de Mestrado Integrado em Engenharia e Gestão Industrial. Foram estudadas as propriedades dos itens e a fiabilidade de três testes com itens de escolha múltipla utilizados no âmbito da avaliação contínua.

Para analisar a qualidade de itens/testes e comparar as classificações obtidas pelos alunos que frequentaram a unidade curricular no ano lectivo 2006/2007 e no ano lectivo 2007/2008, foi considerado cada conjunto de alunos como amostras independentes da mesma população.

A análise dos instrumentos permitiu identificar os itens discriminativos e os níveis de dificuldade dos diversos itens bem como a consistência interna dos instrumentos, paralelamente com a função de informação dos itens e do teste. Apesar de as duas abordagens (TCT e MRI) coincidirem no que diz respeito aos achados principais, adicionalmente foi realçada a importância da análise baseada em MRI, pela possibilidade da criação de bancos de itens e assim, os itens/testes utilizados na aferição das aprendizagens, permitirem a comparação temporal dos resultados [Costa, Oliveira e Ferrão (2008); Ferrão, Costa e Oliveira (2009)].

Referências

- AERA, APA, NCME (1999). *Standards for Educational and Psychological Testing*. American Educational Research Association, American Psychological Association, National Council on Measurement in Education.
- Alarcão, I. (2000). “Para uma Conceitualização dos Fenómenos de Sucesso/Insucesso Escolares no Ensino Superior”. Em J. Tavares e R. Santiago (Orgs.), *Ensino Superior: (In)sucesso Académico*. Porto: Porto Editora.
- Birnbaum, A. (1968). “Some latent trait models and their use in inferring an examinee's ability”. Em F. Lord. e M. Novick (Eds.), *Statistical Theories of Mental Test Scores*. Reading, Mass: Addison-Wesley.
- Cosme, C. I. R. (2004). *Aplicação do Modelo de Regressão Linear no Estudo dos Factores Associados ao Sucesso Escolar nos Cursos do Ensino Superior: Caso UBI*. Dissertação de Mestrado não publicada. Covilhã: Universidade da Beira Interior.
- Costa, P, Oliveira P. e Ferrão, M. E. (2008). “Statistical Issues on Multiple Choice Tests in Engineering Assessment: A Classical Test Theory Approach”. Submetido para publicação.
- Ferrão, M.E., Costa, P. e Oliveira, P.N. (2009). “Item Response Model Applied to Developing a Common Metric: statistics assessment in Engineering”. Working paper.
- Cronbach, L.J. (1951). “Coefficient alpha and the internal structure of tests”. *Psychometrika*, 16, 297-334.
- Decreto-Lei nº 42/2005. D.R. I Série-A 37 (2005-02-22) 1494-1499. Imprensa Nacional Casa da Moeda, Lisboa, Portugal.
- Dunn, G. (1989). *Design and Analysis of Reliability Studies: The Statistical Evaluation of Measurement Errors*. London: Edward Arnold.
- ENQA (2005). *Standards and Guidelines for Quality Assurance in the European Higher Education Area*. European Association for Quality Assurance in Higher Education. Helsinki, Finland.
- Ferrão, M.E. (2008). “E-assessment within the Bologna paradigm: evidence from Portugal”. Submetido para publicação.
- Hambleton, R.K., Swaminathan, H. e Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. North Caroline: Sage Publications.
- Hand, D. J. (2004). *Measurement, Theory and Practice*. London. Arnold.
- Holmes, P. (2002). “Teaching, Learning and Assessment”. RSS Centre for Statistical Education. Online em <http://mathstore.ac.uk/newsletter/aug2000/pdf/tla.pdf> .
- Jolliffe, F. (1997). “Issues in Constructing Assessment Instruments for the Classroom”. Em I. Gal e J. B. Garfield (eds.), *The Assessment Challenge in Statistics Education*. IOS Press (Int. Statist. Institute).
- Lord, F. M. e Novik, M. R. (1968). *Statistical Theory of Mnetal Test Scores*. Reading, M. A. Addison-Wesley.
- Miranda, M. J. e Marques, J. F. (1996). “Sobre o Acesso ao Ensino Superior em Portugal. Resultados e Implicações de um estudo Sectorial”. Revista *Iberoamericana de Diagnóstico y Evaluación Psicológica*, 2, 199-213.
- Nunnally, J.C. e Durham, R.L. (1975). “Validity, reliability, and special problems of measurement in evaluation research”. Em E. L Struening e M. Guttentag (eds.), *Handbook of Evaluation Research*, vol.1. London: Sage.
- Petocz, P. e Reid, A. (2007). “Learning and Assessment in Statistics”. International Statistical Institute, IASE Satellite.
- Soeiro, A e Cabral, J. (2004). “Engineering students assessment at University of Porto”. *European Journal of Engineering Education*, 29, 283 – 290.

Steif, P. e Dantzer, J. (2005). "A Statics Concept Inventory: Development and Psychometric Analysis". *Journal of Engineering Education*, 33, 363-371.

Steif, P. S. e Hansen, M. A. (2006). "Comparisons between performances in a Statics Concept Inventory and Course Examinations". *International Journal of Engineering Education*, 22, 1070-1076.

Tavares, J. Santiago, R. e Lencastre, L. (1998). *Insucesso no Primeiro Ano do Ensino Superior: Um Estudo no Âmbito dos Cursos de Licenciatura em Ciências e Engenharia na Universidade de Aveiro*. Aveiro: TIPAVE.

Teixeira, J., Silva, J. e Flores, P. (2007). "Development of mechanical engineering curricula at the University of Minho". *European Journal of Engineering Education*, 32, 539-549.

Wakeford, R. (2003). "Principles of student assessment". Em H. Fry, S. Ketteridge, e S. A. Marshall, *A Handbook for Teaching & Learning in Higher Education* (2nd ed). London: Routledge Falmer.



A função estatística do Banco de Portugal

João Cadete de Matos, António Garcia, Filipa Lima, José Sérgio Branco
jcmatos@bportugal.pt, agarcia@bportugal.pt, slima@bportugal.pt, jscbranco@bportugal.pt

*Departamento de Estatística
Banco de Portugal*

1. ENQUADRAMENTO E ORGANIZAÇÃO

A função estatística do Banco de Portugal foi desenvolvida como resultado da sua missão histórica de condução da política monetária e execução da política cambial, bem como das suas funções de gestão das disponibilidades externas, de velar pela estabilidade do sistema financeiro e de aconselhar o Governo nos domínios económico e financeiro. Estas missões levaram à necessidade de obtenção de estatísticas fiáveis, nomeadamente nas áreas monetária, financeira, cambial e de balança de pagamentos.

A Lei Orgânica do Banco de Portugal¹ consagra, no seu Artigo 13º, a responsabilidade do Banco de Portugal na “recolha e elaboração das estatísticas monetárias, financeiras, cambiais e da balança de pagamentos, designadamente no âmbito da sua colaboração com o Banco Central Europeu”. Adicionalmente, em 2008 entrou em vigor a nova Lei do Sistema Estatístico Nacional² que, entre outros aspectos, reconhece a qualidade de autoridade estatística ao Banco de Portugal e consagra as atribuições do Banco de Portugal no âmbito do Sistema Estatístico Nacional (SEN), em perfeita consonância com as previstas na sua Lei Orgânica.

Relativamente ao enquadramento internacional, existe uma dupla envolvente. Por um lado, o Banco de Portugal integra o Sistema Europeu de Bancos Centrais, que, de acordo com os seus estatutos³, tem competências próprias de recolha de informação estatística. Por outro lado, o Banco de Portugal integra também o Sistema Estatístico Europeu que compreende, para além do Eurostat e do Banco Central Europeu, os institutos nacionais de estatística e os bancos centrais nacionais da União Europeia. Em ambos os casos, o principal objectivo é a harmonização metodológica no espaço europeu e a definição de requisitos comuns de transmissão de estatísticas.

Em particular nos últimos dez anos, as responsabilidades de produção estatística resultantes dos compromissos assumidos pelo Banco de Portugal, tanto a nível nacional como internacional, registaram um aumento significativo. Para corresponder às crescentes exigências de recursos associados às actividades de produção estatística, em particular as relacionadas com a preparação da União Monetária, foi decidido, em 1997, centralizar a função estatística num único Departamento do Banco de Portugal, o Departamento de Estatística. Este processo permitiu ganhos de eficiência significativos e garantiu uma coerência e harmonização acrescidas na produção das diferentes estatísticas, aspectos determinantes na produção de estatísticas atempadas e de qualidade.

¹ Aprovada pela Lei n.º 5/98, de 31 de Janeiro, com as alterações introduzidas pelos Decretos-Leis n.º 118/2001, de 17 de Abril, n.º 50/2004, de 10 de Março e n.º 39/2007, de 20 de Fevereiro.

² Lei n.º 22/2008, de 13 de Maio.

³ Protocolo anexo ao Tratado que institui a Comunidade Europeia (JO C 191 de 29.7.1992, p. 68), alterado pelo Tratado de Amesterdão (JO C 340 de 10.11.1997, p.1), pelo Tratado de Nice (JO C 80 de 10.3.2001, p. 1), pela Decisão do Conselho 2003/223/CE (JO L 83 de 1.4.2003, p. 66) e pelo Acto de Adesão da República Checa, da Estónia, de Chipre, da Letónia, da Lituânia, da Hungria, de Malta, da Polónia, da Eslovénia e da Eslováquia (JO L 236, 23.9.2003, p. 33).

Actualmente, para além da utilização interna, servindo de base a análises e de suporte a tomada de decisões no âmbito das suas funções, e do reporte aos Organismos Internacionais (Banco Central Europeu, Eurostat, Fundo Monetário Internacional, Banco de Pagamentos Internacionais e Organização para a Cooperação e Desenvolvimento Económico), as estatísticas produzidas pelo Banco de Portugal são disponibilizadas ao público, de forma gratuita e simultânea a todos os utilizadores, no sítio do Banco de Portugal na Internet: Boletim Estatístico do Banco de Portugal⁴ e no BPstat – Estatísticas *online* – serviço de disseminação estatística na Internet⁵.

O BPstat está disponível a partir da página principal do sítio do Banco de Portugal na Internet, permitindo acesso *online* à Base de Dados Estatística Interactiva do Banco de Portugal. O seu principal objectivo consiste em facultar um acesso fácil e célere a um vasto conjunto de séries estatísticas e indicadores produzidos pelo Banco de Portugal e por outras instituições, bem como às estatísticas que integram a divulgação conjunta do Eurosistema. Este serviço está disponível em duas versões, respectivamente em português e em inglês.

O Banco de Portugal mantém elevados padrões de qualidade na produção e disseminação estatística, factor determinante para consolidar a confiança que os utilizadores depositam nas estatísticas da sua responsabilidade. Este mesmo entendimento está consagrado no Compromisso Público do SEBC relativamente à sua função estatística (Maio de 2007)⁶, a qual é orientada pelos princípios da imparcialidade, independência científica, confidencialidade estatística, custos apropriados e parcimónia nos pedidos de informação junto dos reportantes e, ainda, elevada qualidade na produção e disseminação estatística

2. DOMÍNIOS ESTATÍSTICOS

O Banco de Portugal recolhe e elabora estatísticas em vários domínios. As sub-secções seguintes descrevem brevemente cada um deles, a saber,

- Estatísticas das Instituições Financeiras Monetárias,
- Estatísticas das Instituições Financeiras Não Monetárias,
- Estatísticas da Central de Responsabilidades de Crédito,
- Estatísticas de Mercados de Títulos,
- Estatísticas da Balança de Pagamentos,
- Estatísticas da Posição de Investimento Internacional,
- Estatísticas Cambiais,
- Estatísticas das Finanças Públicas,
- Estatísticas das Contas Nacionais Financeiras,
- Estatísticas das Empresas Não Financeiras da Central de Balanços
- Estatísticas de Sistemas de Pagamentos,
- Estatísticas de Emissão Monetária.

Para uma apresentação mais detalhada destas estatísticas podem consultar-se os respectivos documentos metodológicos e suplementos ao Boletim Estatístico, disponíveis no sítio do Banco de Portugal na Internet⁷, assim como a meta informação disponibilizada no BPstat – Estatísticas *online*. O

⁴ Boletim Estatístico do Banco de Portugal (também disponível em papel e em CD-ROM), com periodicidade mensal, sendo possível obter, a partir das versões electrónicas, ficheiros de formato csv com dados e com meta informação. São também publicados suplementos a este boletim e notas de informação estatística.

⁵ O BPstat pode ser acedido através de <http://apl1.bportugal.pt/estatisticasweb>.

⁶ O Compromisso público quanto à função estatística do BCE pode ser acedido através de <http://www.ecb.int/stats/html/pcstats.pt.html>

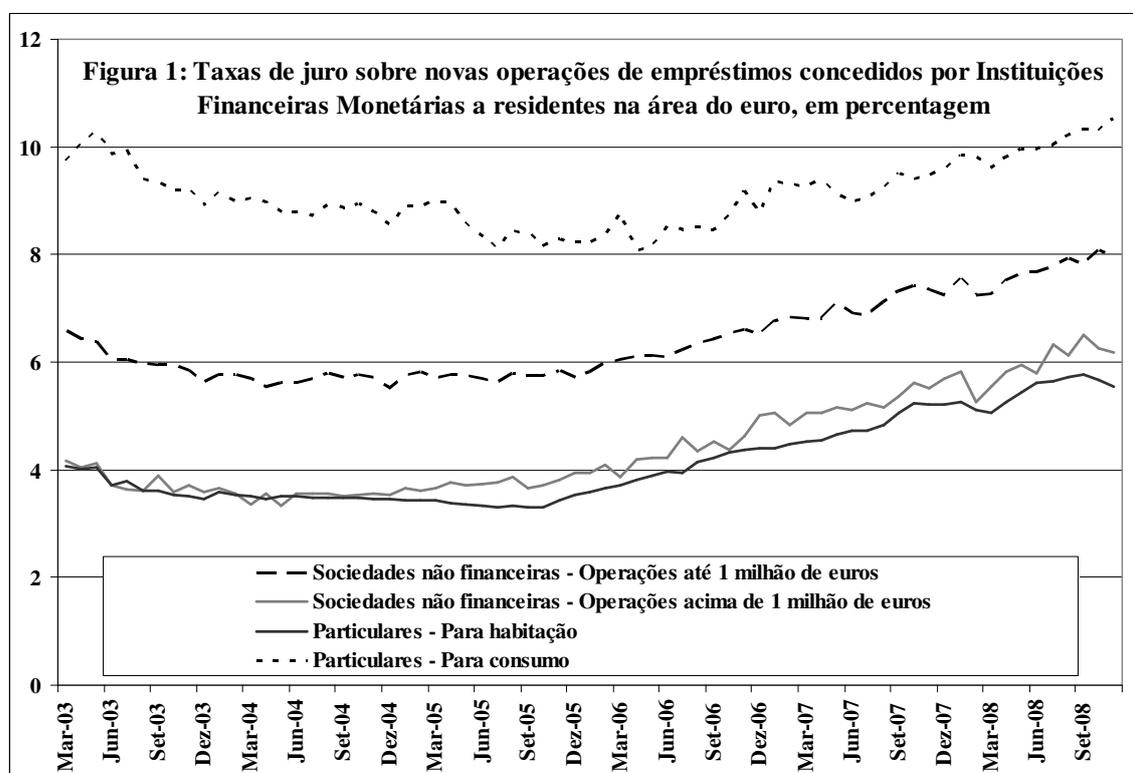
⁷ <http://www.bportugal.pt>

plano de actividade estatística do Banco de Portugal para 2009, também disponível na Internet⁸, permite ter uma ideia mais clara do conjunto de actividades desenvolvidas pelo Banco de Portugal nesta matéria.

2.1. Estatísticas das Instituições Financeiras Monetárias

As estatísticas de balanço das instituições financeiras monetárias (IFM) abrangem os agregados monetários e respectivas contrapartidas internas e externas, a síntese monetária, o balanço do Banco de Portugal e o balanço agregado e os custos e proveitos por natureza das Outras IFM que integram, no caso de Portugal, as seguintes entidades: bancos, caixas de crédito agrícola mútuo, caixas económicas e os fundos de mercado monetário.

O Banco de Portugal elabora também estatísticas de taxas de juro das IFM, com base em metodologias que recorrem à amostragem não aleatória estratificada de acordo com procedimentos harmonizados a nível europeu. Estas estatísticas cobrem as operações de empréstimos e depósitos denominados em euros face a particulares e sociedades não financeiras residentes em países pertencentes à União Monetária, com informação de taxas de juro sobre novas operações e de saldos, e respectivos montantes associados.



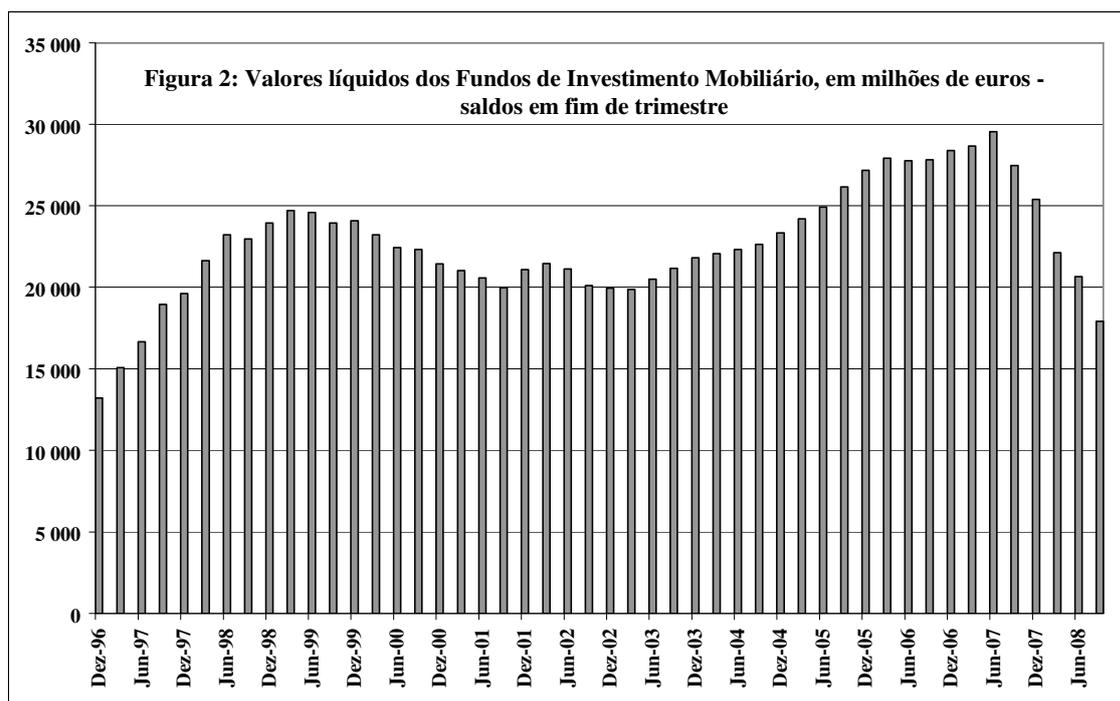
2.2. Estatísticas das Instituições Financeiras Não Monetárias

É também recolhida informação sobre estatísticas de instituições financeiras não monetárias (IFNM) excepto sociedades de seguros e fundos de pensões (SSFP). As instituições financeiras que integram actualmente o sector das IFNM excepto SSFP são: (i) os outros intermediários financeiros⁹, cuja função principal consiste em fornecer serviços de intermediação financeira (não monetária), isto é, em adquirir activos e incorrer, por sua própria conta, em passivos sob outras formas que não numerário, depósitos e/ou substitutos próximos destes junto de unidades institucionais que não IFM, e (ii) os

⁸ http://cse.ine.pt/ngt_server/attachfileu.jsp?look_parentBoui=53355743&att_display=n&att_download=y

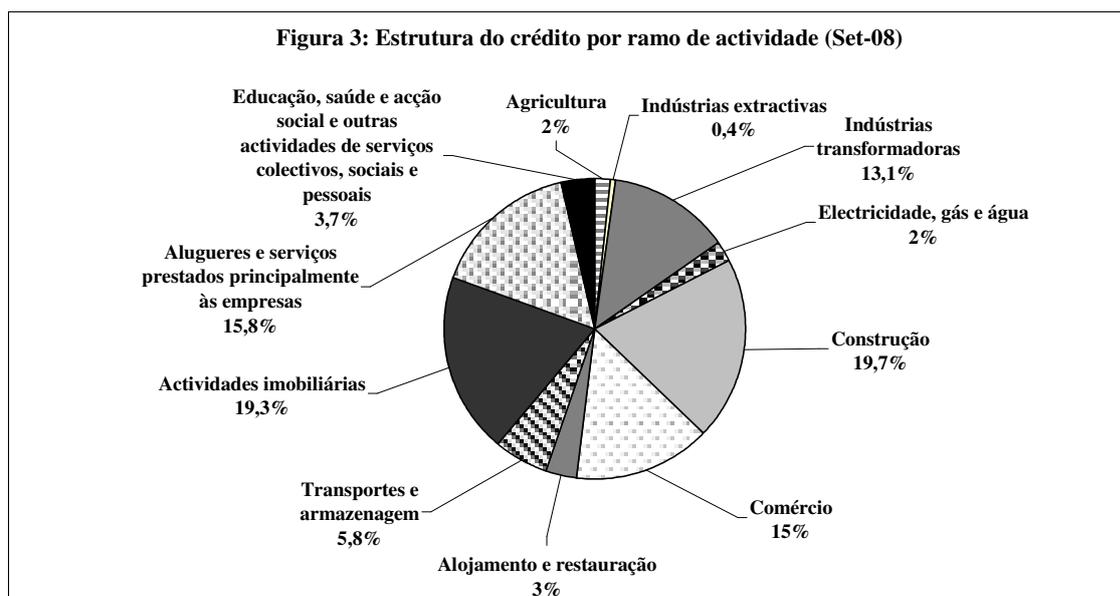
⁹ A título exemplificativo: Fundos de Investimento (Mobiliários e Imobiliários), Fundos e Sociedades de Capital de Risco, Fundos e Sociedades de Titularização De Créditos, Sociedades de Factoring, Sociedades de Locação Financeira, Sociedades Financeiras para Aquisições a Crédito, Sociedades Financeiras de Corretagem, Sociedades Gestoras de Participações Sociais (do Sector Financeiro), Sociedades de Garantia Mútua, Instituições Financeiras de Crédito.

auxiliares financeiros¹⁰ cuja principal função consiste em exercer actividades financeiras auxiliares, isto é, actividades estritamente ligadas à intermediação financeira sem contudo fazerem parte dela.



2.3. Estatísticas da Central de Responsabilidades de Crédito

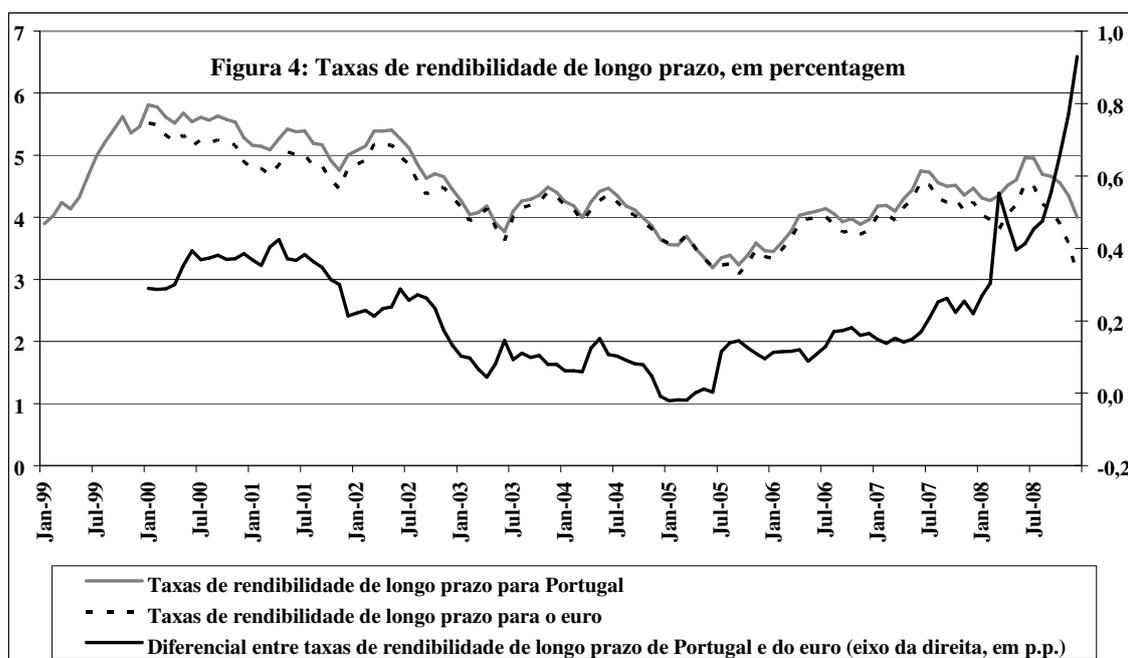
É publicada informação estatística trimestral sobre créditos concedidos por instituições financeiras a sociedades não financeiras com base na informação da Central de Responsabilidades de Crédito, uma base de dados gerida pelo Banco de Portugal. Os resultados da exploração estatística desta base de dados permitem obter informação relativa à repartição territorial, desagregação por ramo de actividade económica e distribuição não só do crédito concedido como também do crédito vencido.



¹⁰ A título exemplificativo: Agências de Câmbios, Auxiliares de Seguros, Sociedades Corretoras, Sociedades Gestoras de Fundos De Investimento, Sociedades Gestoras de Fundos de Pensões, Sociedades Gestoras de Patrimónios.

2.4. Estatísticas de Mercados de Títulos

O Banco de Portugal divulga informação mensal sobre emissões de títulos (ações e títulos de dívida) emitidos por residentes em Portugal. Esta informação está organizada em dois grandes blocos, um para transacções, que englobam emissões brutas, amortizações / cessações e emissões líquidas, e outro para posições em fim de período. Ambos os blocos apresentam informação desagregada por tipo de título, prazo, moeda e mercado de emissão. Estes blocos, para além de conterem séries relativas ao total da economia, subdividem-se ainda em quatro conjuntos de séries, de acordo com o sector institucional da entidade emitente dos títulos. As séries relativas às emissões do Sector Não Financeiro encontram-se, ainda, desagregadas por ramo de actividade económica da entidade emitente dos títulos.



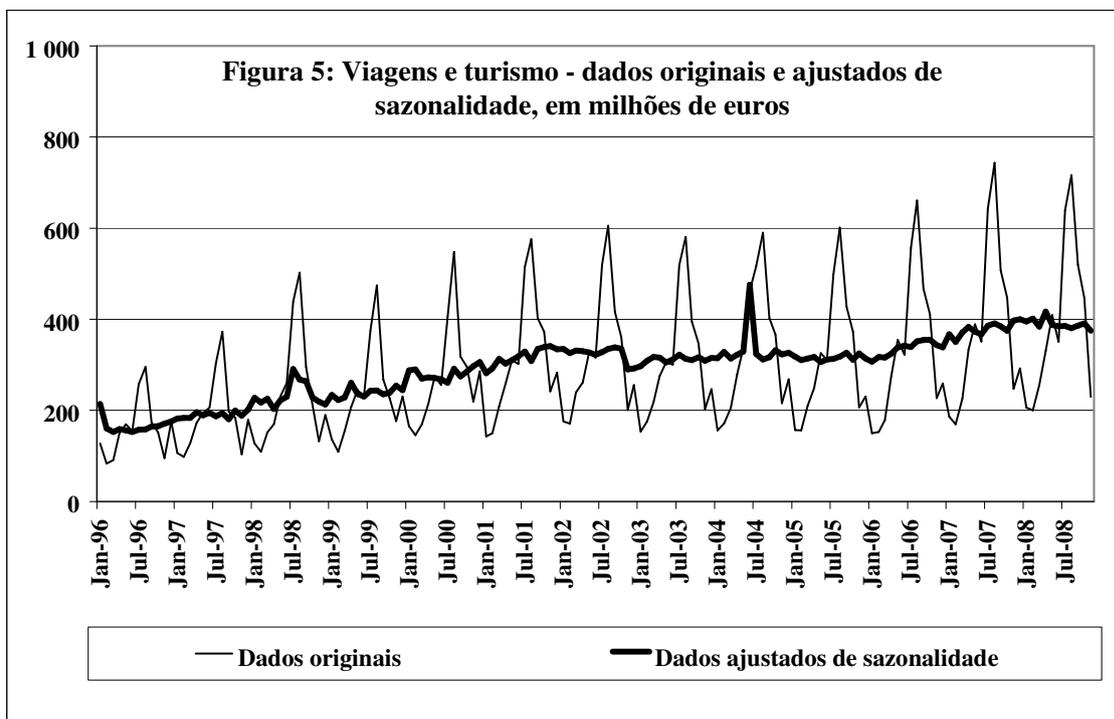
2.5. Estatísticas da Balança de Pagamentos

As estatísticas da balança de pagamentos registam de forma sistemática, e para um determinado período, o conjunto das transacções económicas de uma economia face ao resto do Mundo (o exterior). A balança de pagamentos é composta pelas balanças corrente, de capital e financeira.

A balança corrente inclui as transacções em bens e serviços (exportações e importações), os fluxos financeiros relacionados com rendimentos da propriedade (juros e dividendos pagos) e do trabalho e outras transferências correntes (por exemplo, da União Europeia e remessas de emigrantes).

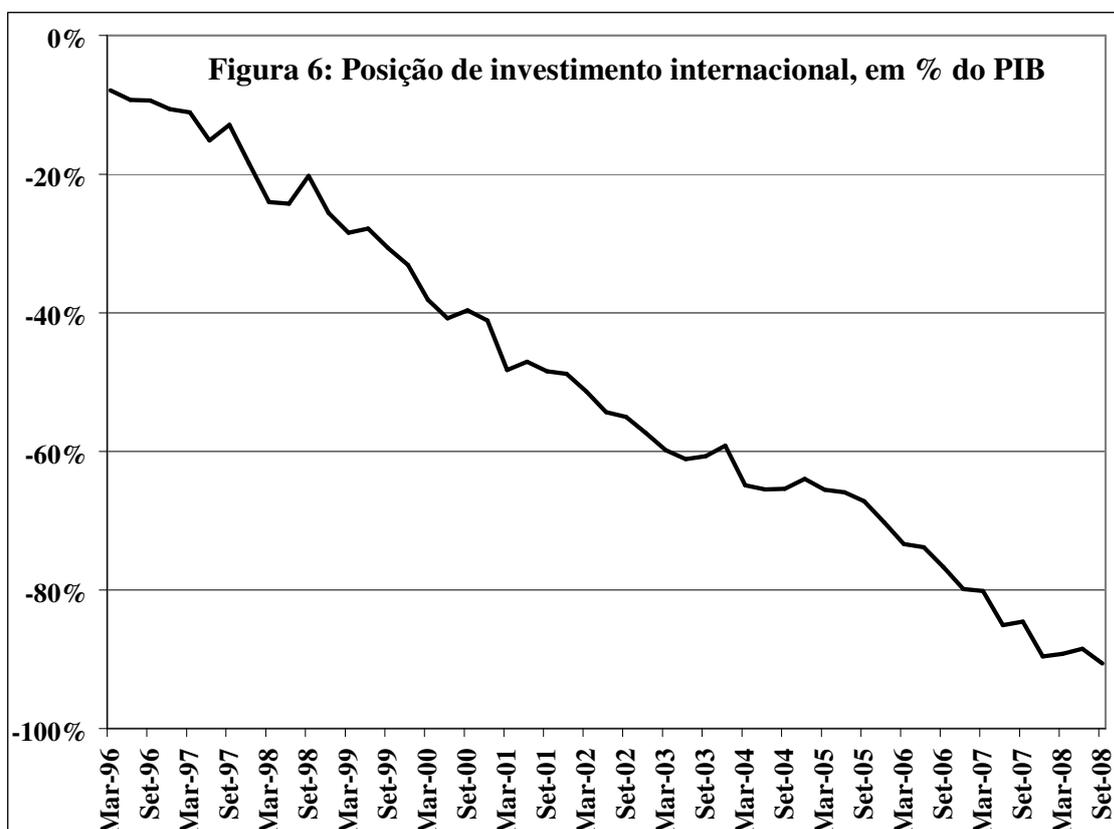
A balança de capital consiste fundamentalmente em transferências internacionais destinadas a investimento ou a mudanças de propriedade sem contrapartida. O conjunto das balanças corrente e de capital constitui um agregado da maior importância para a análise da situação económica de um país, pois permite conhecer a situação da economia nacional face a não residentes.

A balança financeira constitui a terceira componente da balança de pagamentos, sendo o reflexo, ao nível das transacções em instrumentos financeiros, das operações correntes e de capital. A balança financeira apresenta a forma de financiamento da economia face ao exterior, com desagregação por instrumento financeiro, por sector institucional e por país de contrapartida. As transacções em investimento directo estrangeiro em Portugal e em investimento directo de Portugal no estrangeiro são um dos resultados mais relevantes desta componente da balança de pagamentos.



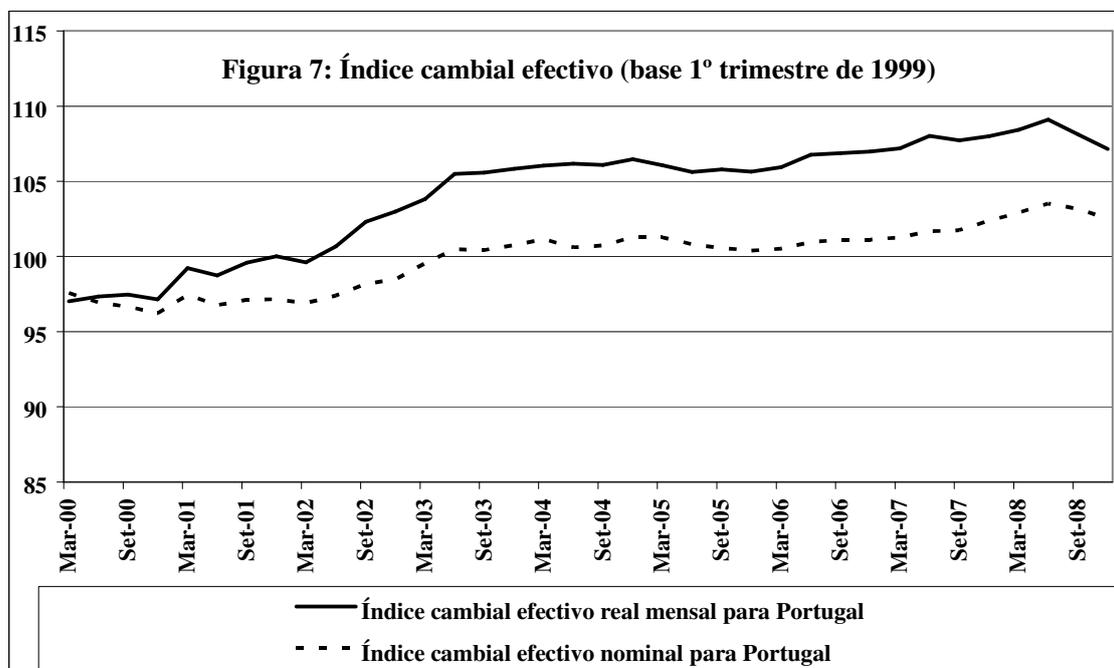
2.6. Estatísticas da Posição de Investimento Internacional

A posição de investimento internacional corresponde aos *stocks* no final de cada trimestre dos activos financeiros e dos passivos de Portugal face ao exterior. As diferenças entre posições de períodos consecutivos reflectem transacções financeiras (registadas na balança de pagamentos), variações de preço, variações cambiais e outros ajustamentos, que afectam o nível de activos e/ou passivos externos no final do período.



2.7. Estatísticas Cambiais

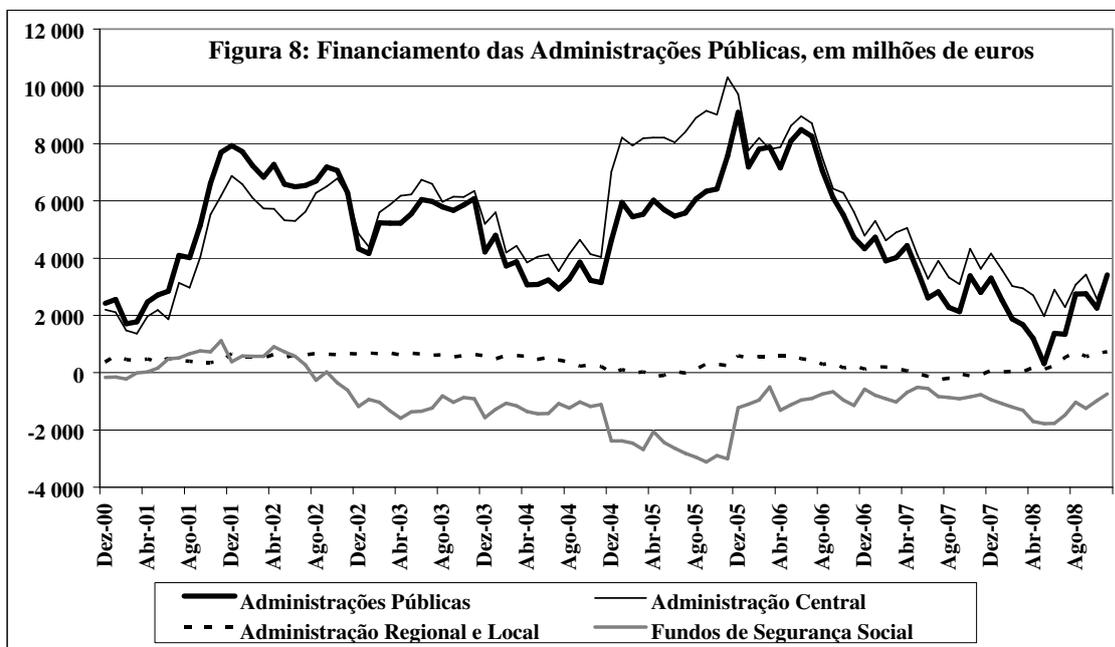
O Banco de Portugal elabora e publica estatísticas sobre as taxas de câmbio do euro face a outras moedas. São apresentados os valores mensais médios e em fim de mês. São também apurados índices de taxa de câmbio efectiva nominal e real. Os índices cambiais efectivos nominais são uma medida de síntese do valor externo de uma moeda em relação às moedas dos principais parceiros comerciais. As taxas de câmbio efectivas reais (obtidas deflacionando as taxas nominais com índices adequados de preços e custos) constituem os indicadores mais utilizados relativos à competitividade externa da economia. Finalmente, são disponibilizados dados sobre o movimento do mercado de câmbios.



2.8. Estatísticas das Finanças Públicas

Com base nas estatísticas monetárias e financeiras e nas estatísticas da balança de pagamentos, são elaboradas mensalmente estatísticas do financiamento das administrações públicas. Este conjunto de informação permite acompanhar ao longo do ano a evolução da situação financeira das administrações públicas. A informação é disponibilizada de forma desagregada para cada subsector das administrações públicas (administração central, administração regional e local e fundos de segurança social). São também apresentados quadros estatísticos sobre a dívida directa do Estado e sobre a execução orçamental do Estado e da Segurança Social, com base em informação do Ministério das Finanças.

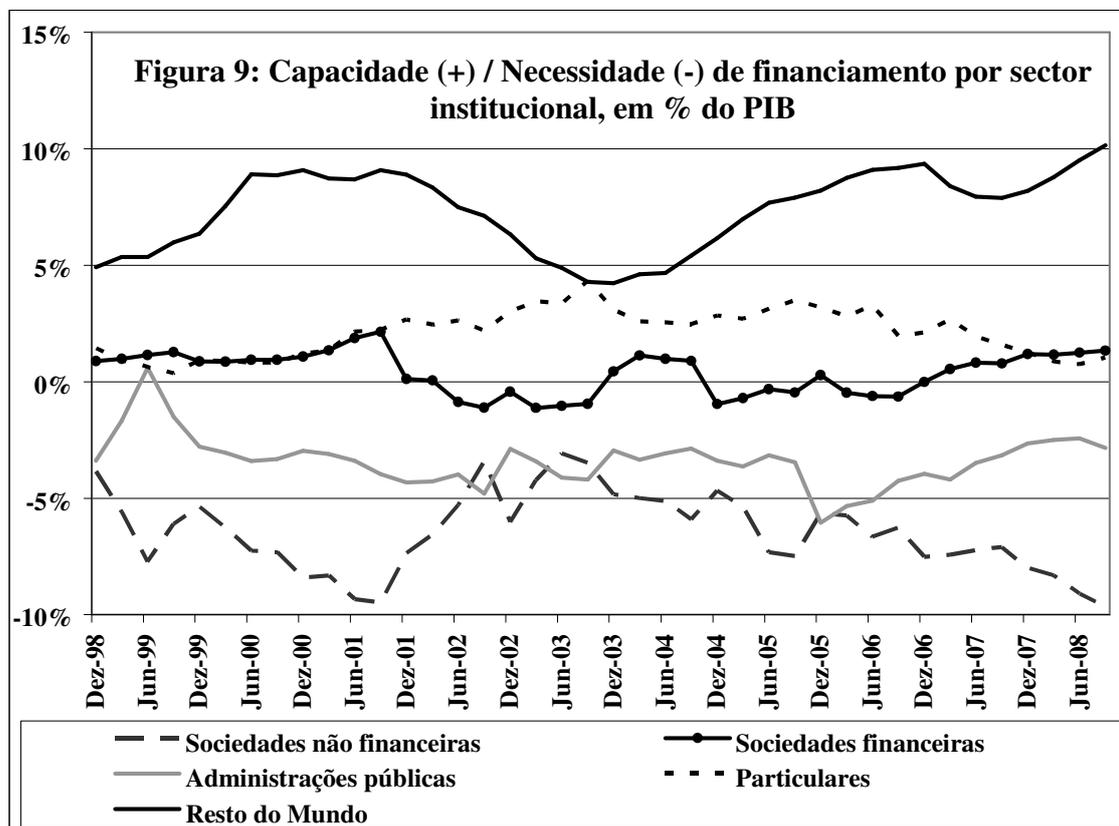
Adicionalmente, o Banco de Portugal colabora com o INE e com o Ministério das Finanças na elaboração das notificações de Portugal no âmbito do Procedimento dos Défices Excessivos. Este exercício tem como objectivo o apuramento dos valores anuais do défice e da dívida bruta das administrações públicas para efeitos de avaliação, por parte da Comissão Europeia, da situação orçamental portuguesa. Os dados são compilados com base em informação de contas nacionais financeiras e não financeiras.



2.9. Estatísticas das Contas Nacionais Financeiras

Estas estatísticas são uma componente das contas nacionais portuguesas e permitem analisar a informação estatística relativa às transacções e patrimónios financeiros da economia, com desagregação por sector institucional (Sociedades Não Financeiras, Sociedades Financeiras, Administrações Públicas, Particulares – cobrindo as Famílias e as Instituições Sem Fins Lucrativos ao Serviço das Famílias – e Resto do Mundo) e por instrumento financeiro. A periodicidade das contas nacionais financeiras é trimestral. As contas nacionais são elaboradas com base na metodologia do Sistema Europeu de Contas Nacionais e Regionais, diploma legislativo que harmoniza a elaboração de contas nacionais nos países da União Europeia. Este sistema permite a obtenção de estatísticas integradas entre os vários sectores institucionais e com as outras componentes das contas nacionais. Esta informação é compilada com base em dados disponibilizados pelas outras áreas de estatísticas do Banco de Portugal, bem como em informação diversa recolhida junto de outros organismos (Instituto de Seguros de Portugal e Ministério das Finanças, por exemplo).

A representação estatística sintética da estrutura de financiamento da economia, que as Contas Financeiras facultam, é possível porque estas são construídas num formato matricial, com informação organizada de modo a identificar os agentes económicos desagregados por sectores institucionais, e as operações financeiras desagregadas por instrumento financeiro. Os principais resultados destas estatísticas são a capacidade ou necessidade de financiamento e a riqueza financeira de cada sector institucional e do país.

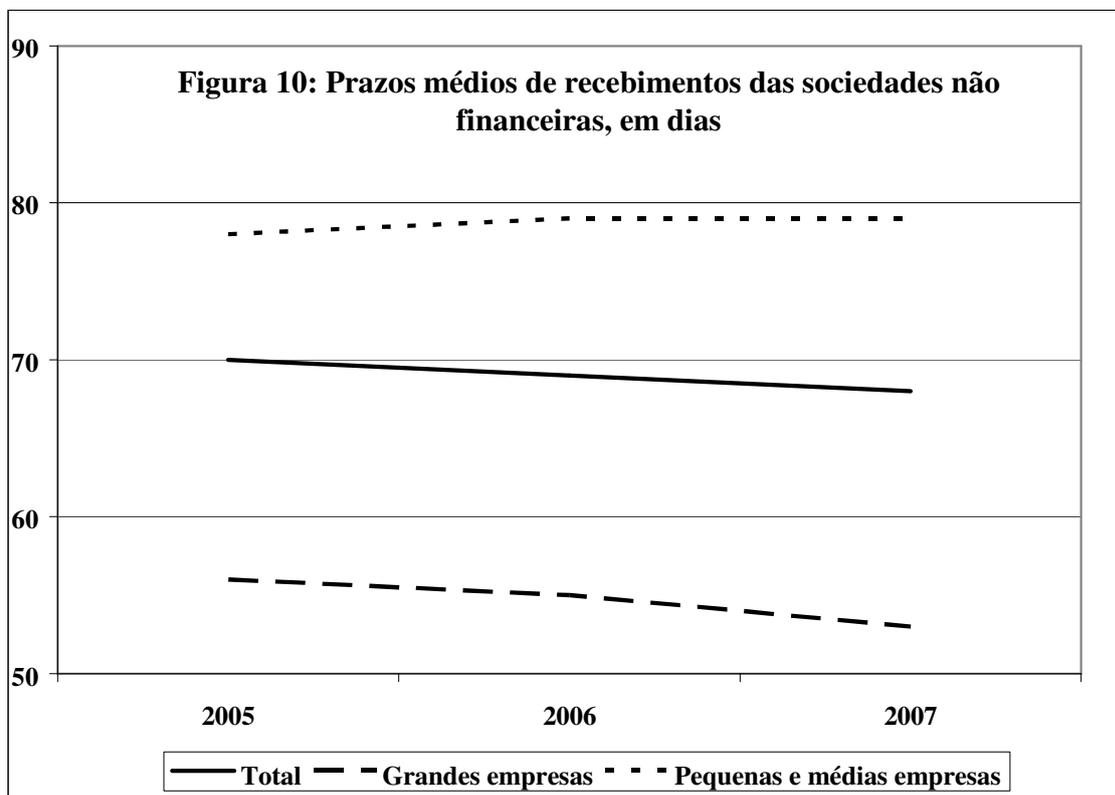


2.10. Estatísticas das Empresas Não Financeiras da Central de Balanços

O Banco de Portugal elabora e publica um conjunto de informação sobre empresas não financeiras, com base em informação contabilística disponibilizada por estas entidades. Até 2006, dispunha-se de uma amostra de empresas que efectuavam o envio voluntário da sua informação contabilística. A partir de 2007, com a implementação da Informação Empresarial Simplificada (IES), passou a dispor-se de uma cobertura quase total do universo de empresas não financeiras. A IES é uma iniciativa conjunta do Ministério da Justiça, Ministério das Finanças, INE e Banco de Portugal, que permitiu que as empresas passassem a enviar apenas um conjunto de informação, em formato electrónico, em lugar dos anteriores quatro envios de informação para as quatro entidades.

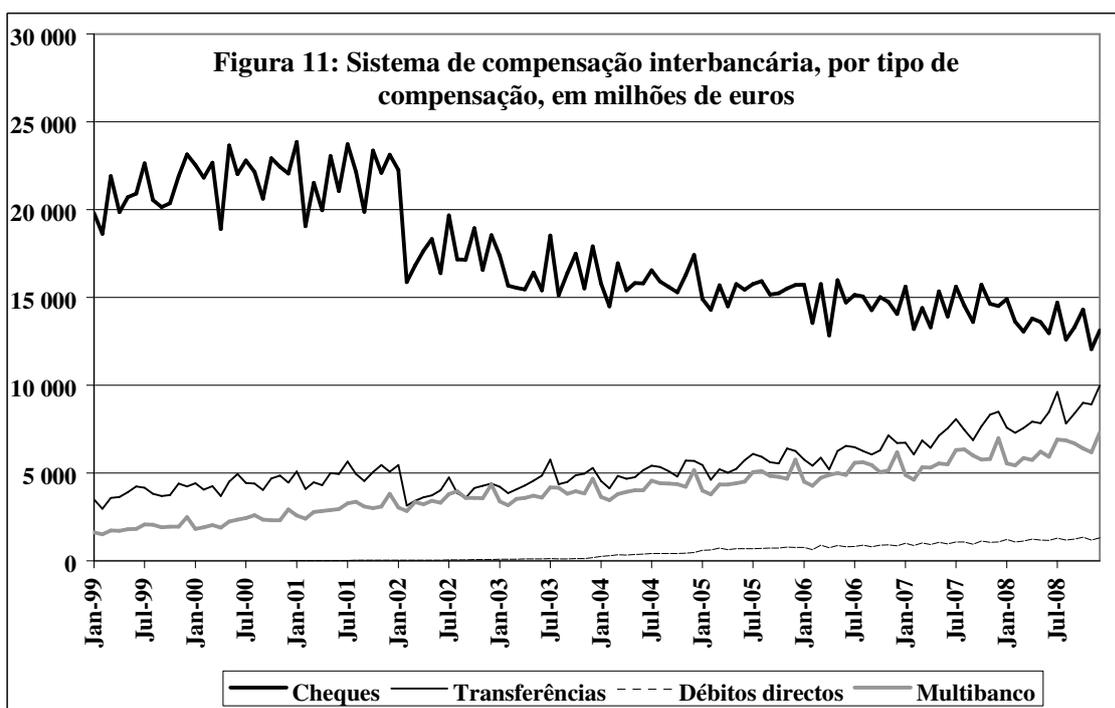
Para além da cooperação relativa à informação anual, o Banco de Portugal e o INE colaboram também no Inquérito Trimestral às Empresas não Financeiras dirigido a uma amostra de empresas. Este instrumento permite obter informação de frequência infra-anual para as sociedades não financeiras.

Os principais resultados divulgados são estruturas de balanços e demonstrações de resultados, por ramo de actividade e classes de dimensão das empresas, bem como um conjunto de indicadores sob a forma de rácios e de taxas de variação de agregados económicos e financeiros. Para além da informação disponibilizada no Boletim Estatístico, é também divulgado na página da Internet do Banco de Portugal um conjunto de indicadores com grande detalhe a nível sectorial (“Quadros do Sector”). Este tipo de informação integra também bases de dados internacionais com informação comparável sobre outros países (Base de Dados BACH – *Bank for the Accounts of Companies Harmonised* e Base de Dados Referências Sectoriais Europeias).



2.11. Estatísticas de Sistemas de Pagamentos

O Banco de Portugal disponibiliza anualmente um conjunto de informação detalhada sobre sistemas de pagamentos, nomeadamente no “Relatório dos Sistemas de Pagamentos e de Liquidação Interbancária” e na publicação “Sistemas de Pagamento em Portugal”¹¹, disponíveis na página da Internet do Banco, bem como divulga mensalmente no Boletim Estatístico um conjunto de informação sobre o sistema de compensação interbancária (por exemplo movimentos de cheques, transferências interbancárias, débitos directos e operações no Multibanco) e sobre o sistema de pagamentos de grandes transacções (operações domésticas e transnacionais).

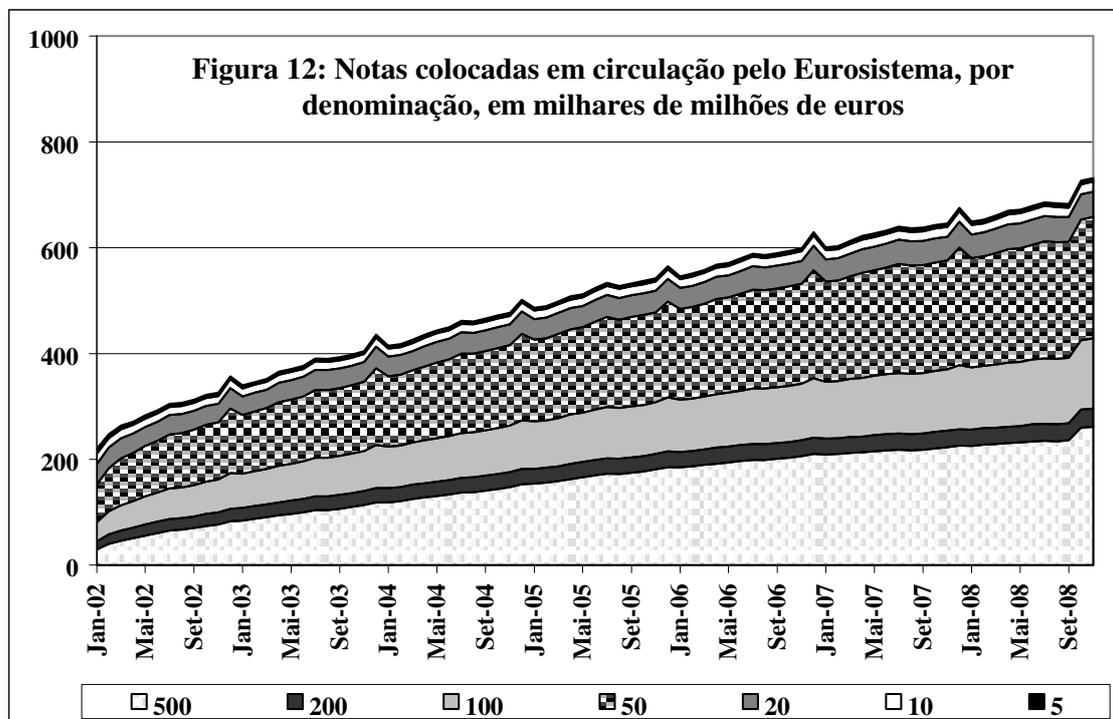


¹¹ Corresponde ao capítulo sobre Portugal Payment Systems/Payments & Markets/Blue Book, cuja informação estatística está disponível em

<http://sdw.ecb.europa.eu/reports.do?node=100000760>

2.12. Estatísticas de Emissão Monetária

O Banco de Portugal disponibiliza informação estatística sobre a emissão de notas e moedas, nomeadamente sobre o montante em euros colocado em circulação pelo Banco de Portugal e pelo Eurosistema. É igualmente disponibilizada informação relativa a depósitos e levantamentos de notas e moedas realizados no Banco de Portugal, assim como informação sobre notas contrafeitas detectadas em Portugal.



3. DESAFIOS PARA O FUTURO

É entendimento do Banco de Portugal, que a prossecução de elevados padrões de qualidade no âmbito da actividade de produção e disseminação estatística constitui um factor determinante para consolidar a confiança que os utilizadores depositam nas estatísticas da sua responsabilidade. Assim, o Banco de Portugal continuará a ter como principais prioridades, no exercício da sua função estatística, a implementação de métodos eficientes de recolha, compilação e difusão de informação.

Simultaneamente, a crescente integração a nível europeu e global conduz a uma cada vez maior exigência dos organismos internacionais ao nível de quantidade e qualidade da informação transmitida pelos países. A constante evolução das economias e dos mercados financeiros obriga também a uma monitorização constante da adaptação dos padrões estatísticos à realidade.

Em 2008 prosseguiram os trabalhos de revisão do Sistema de Contas Nacionais das Nações Unidas e do Manual de Balança de Pagamentos e da Posição de Investimento Internacional do Fundo Monetário Internacional, tendo sido também iniciados os trabalhos de revisão do Sistema Europeu de Contas. A revisão em simultâneo destes manuais visa assegurar a consistência conceptual e metodológica entre eles. As alterações introduzidas respeitam essencialmente à adaptação dos manuais internacionais de compilação de estatísticas aos novos fenómenos que caracterizam as economias cada vez mais sofisticadas, globalizadas e interligadas, com necessidades de informação cada vez mais exigentes. O Banco de Portugal tem acompanhado o processo e intervido activamente, apresentando comentários e propostas de alteração

A adaptação dos sistemas de recolha de informação e de produção estatística aos novos manuais internacionais de compilação de estatísticas, continuando a assegurar elevados padrões de qualidade, constitui, assim, um dos próximos desafios para a função estatística do Banco de Portugal.

Referências

Boletim Estatístico do Banco de Portugal (http://www.bportugal.pt/publish/statbol/html/be_pub_p.htm)

Suplementos ao Boletim Estatístico do Banco de Portugal:

(http://www.bportugal.pt/publish/statbol/html/be_pub_p.htm#Suplementos)

2/2008 *Estatísticas de Títulos: Caracterização do Sistema Integrado e Apresentação de Resultados*

1/2008 *Reporte simplificado: incorporação da Informação Empresarial Simplificada nas Estatísticas das Empresas Não Financeiras da Central de Balanços*

1/2007 *Papers presented by Banco de Portugal representatives at the 56th Session of the International Statistical Institute, held in Lisbon, 22 – 29 August 2007*

5/2005 *Estatísticas das Empresas Não Financeiras da Central de Balanços*

4/2005 *Ajustamento Sazonal de Séries Estatísticas da Balança de Pagamentos*

3/2005 *Contas Nacionais Financeiras da Economia Portuguesa. Estatísticas sobre Patrimónios Financeiros de 1999 a 2004*

2/2005 *Contas Nacionais Financeiras da Economia Portuguesa. Notas Metodológicas e Apresentação dos Resultados Estatísticos de 2000 a 2004*

1/2005 *Utilização da Central de Responsabilidades de Crédito no âmbito das Estatísticas Monetárias e Financeiras*

1/2001 *"Balanço estatístico" e "Balanço contabilístico" das Outras Instituições Financeiras Monetárias*

1/2000 *Investimento Directo de Portugal no Exterior*

2/1999 *Informação Estatística sobre Fundos de Investimento Mobiliário (FIM)*

1/1999 *Nova Apresentação das Estatísticas da Balança de Pagamentos*

2/1998 *Investimento Directo do Exterior em Portugal*

1/1998 *Informação Estatística sobre Instituições Financeiras Não Monetárias*

Documentos metodológicos: (http://www.bportugal.pt/stats/Documentos%20Metodologicos/doc_met_p.htm)

Estatísticas de Balanço das Instituições Financeiras Monetárias

Estatísticas de Taxas de Juro das Instituições Financeiras Monetárias

Estatísticas de Instituições Financeiras Não Monetárias

Estatísticas de Emissões de Títulos

Estatísticas de Sistemas de Pagamentos

Estatísticas de Balança de Pagamentos

Estatísticas de Posição de Investimento Internacional

Estatísticas Bancárias Internacionais em Base Consolidada

Estatísticas Cambiais

Contas Nacionais Financeiras

Estatísticas das Sociedades Não Financeiras da Central de Balanços



• Artigos Científicos Publicados

- Afonso, A., Alpizar-Jara, R., Moço, G., Granados, J. E., Pérez, J. M. (2008). Dimensión de la muestra y estimación de su varianza en muestreo por distancias: una aplicación en la estima de abundancia de poblaciones de ungulados de montaña. En *Tendencias actuales en el estudio y conservación de los Caprinos Europeos*. (Granados Torres, J.E., Cano-Manuel, J., Fandos Paris, P. & Cadenas del Llano, R. (Eds.) Junta de Andalucía, Consejería de Medio Ambiente. España. p. 15-29.
- Alpizar-Jara, R. & C.E. Smith (2008). A continuous-time version and a generalization of a Markov-recapture model for trapping experiments. *Mathematical Biosciences*, 214(1-2):11-19.
- Amaral Santos, José A. and Neves, M. Manuela (2008). A Local Maximum Likelihood Estimator for Poisson Regression. *Metrika* 68(3), 257-270. DOI: 10.1007/s00184-007-0156-1, October 2007.
- Antunes, M. e Sousa, L. (2008). Bayesian classification and non-Bayesian label estimation via EM algorithm to identify differentially expressed genes: a comparative study. *Biometrical Journal* 50(5): 824–836. (DOI: 10.1002/bimj.200710468)
- Antunes, N.; Jacinto, G.; e Pacheco, A. (2008). On the minimum hop count and connectivity in one-dimensional ad hoc wireless networks. *Telecommunication Systems* 39, pp. 137–143.
- Boente, G.; Pires, A. M.; e Rodrigues, I. M. (2008). Estimators for the common principal components model based on reweighting: influence functions and Monte Carlo study. *Metrika* 67(2), pp. 189-218.
- Didier, G. e Pipiras, V. (2008). Gaussian stationary processes: adaptive wavelet decompositions, discrete approximations, and their convergence. *J. Fourier Anal. Appl.* 14(2), pp. 203-234.
- Estepa, A.; Estepa, R.; e Pacheco, A. (2008). Accurate resource estimation for generalized VoIP sources. *Telecommunication Systems* 39(1), pp. 37-50.
- Ferreira, F. e Pacheco, A. (2008). Analysis of $GI^X/M(n)/N$ systems with stochastic customer acceptance policy. *Queueing Systems. Theory and Applications* 58(1), pp. 29-55.
- Fraga Alves, I., de Haan, L. and Neves, C. (2009). A test procedure for detecting super heavy tails. *Journal of Statistical Planning and Inference*, 139, pp. 213-227.
- Gomes, M.I., Fraga Alves, M.I. and Araújo Santos, P. (2008). PORT Hill and Moment Estimators for Heavy-Tailed Models. *Commun. Statist. — Simul. & Comput.* 37, 1281-1306, 2008. DOI: 10.1080/03610910802050910, 2008.
- Gomes, M.I. and Henriques Rodrigues, L. (2008). Tail index estimation for heavy tails: accommodation of bias in the excesses over a high threshold. *Extremes* 11(3), 303-328. DOI: 10.1007/s10687-008-0059-1, 2008.
- de Haan, L., Neves, C. and Peng, L. (2008). Parametric tail copula estimation and model testing. *J. Multivariate Analysis* 99:6, 1260-1275.
- Monteiro, J.F., Alpizar-Jara, R., Serrano, E., Crampe, J.P., Pérez, J.M. (2008). Estimación Bayesiana de g_0 usando el muestreo por distancias y su aplicación en las estimas de densidad de ungulados de montaña (*Rupicapra pyrenaica*). En *Tendencias actuales en el estudio y conservación de los Caprinos Europeos*. (Granados Torres, J.E., Cano-Manuel, J., Fandos Paris, P. & Cadenas del Llano, R. (Eds.) Junta de Andalucía, Consejería de Medio Ambiente. España. p. 207-217.
- Monteiro, M. Pereira, I. and Scotto, M. G. (2008). Optimal alarm systems for count processes. *Commun. Statist. Theor. Meth.* 37, 3054-3076.
- Morais, M.C.; Okhrin, Y.; Pacheco, A.; e Schmid, W. (2008). EWMA charts for multivariate output: Some stochastic ordering results. *Communications in Statistics - Theory and Methods* 37(16), pp. 2653-2663.
- Pacheco, A. e Ribeiro, H. (2008). Consecutive customer losses in oscillating $GI^X/M/n$ systems with state dependent services rates. *Annals of Operations Research* 162(1), pp. 143-158.
- Pacheco, A. e Ribeiro, H. (2008). Consecutive customer losses in regular and oscillating $M^X/G/1/n$ systems. *Queueing Systems. Theory and Applications* 58(2), pp.121-136.
- Pacheco, A. e Ribeiro, H. (2008). Moments of the duration of busy periods of $M^X/G/1/n$ systems. *Probability in the Engineering and Informational Sciences* 22(3), pp. 347-354.

- Pérez, J.J. & Alpizar-Jara, R. (2008). Towards the use of distance sampling to monitorize mountain ungulates numbers. In: “*Ecosystem Ecology Research Developments Trends*” (J. Chen & C. Guô, eds.) Nova Science Publishers, Inc., New York. p. 9-20.
- Pipiras, V. e Taqqu, M. S. (2008). Identification of periodic and cyclic fractional stable motions. *Ann. Inst. Henri Poincaré Probab. Stat.* 44(4), pp. 612-637.
- Pipiras, V. e Taqqu, M. S. (2008). Small and large scale asymptotics of some Lévy stochastic integrals. *Methodol. Comput. Appl. Probab.* 10(2), pp. 299-314.
- Ramos, A. and Ledford, A. (2009). A new class of models for bivariate joint tails. *J. R. Statist. Soc., B*, 71 (1). 219-241.
- Sepúlveda, N., Paulino, C. D. e Penha-Gonçalves, C. (2009). Bayesian analysis of allelic penetrance models for complex binary traits. *Computational Statistics and Data Analysis* 53, 1271-1283.
- Serrano, E., Alpizar-Jara, R., Morellet, N., & Hewison, J.M. (2008). A half a century of measuring ungulate body condition using indices. Is it time for a change? *European Journal of Wildlife Research*, 54:675-680.
- Silva, N., Pereira, I. e Silva, M.E. (2008). Estimation and forecasting in SUINAR(1) model. *REVStat*, 6, 253-277.
- Souto-Carneiro, M. M., Fritsch, R., Sepúlveda, N., Lagareiro, M. J., Morgado, N., Longo, N. S., Lipsky, P. E. (2008). The NF-kappaB canonical pathway is involved in the control of the exonucleolytic processing of coding ends during V(D)J recombination. *J Immunol.* 180(2):1040-9.
- Teixeira, A., Rosa, A. e Calapez, T. (2009). Statistical Power Analysis with Microsoft Excel: Normal Tests for One or Two Means as a Prelude to Using Non-Central Distributions to Calculate Power. *Journal of Statistics Education* 17(1).

• Capítulos em Livros

- Gomes, M.I., Pestana, D., Henriques-Rodrigues, L. and Viseu, C. (2008). Tail Behaviour: an Empirical Study. In Arnold, B.C., Balakrishnan, N., Sarabia, J.M. and Mínguez, R. (eds.), *Advances in Mathematical and Statistical Modeling*, Chapter 14, 195-207. ISBN 978-0-8176-4625-7, DOI: 10.1007/978-0-8176-4626-4_14, Birkhauser, Boston.
- Neves, C. and Fraga Alves, M.I. (2008). Ratio of Maximum to the Sum for Testing Super Heavy Tails. In Arnold, B.C., Balakrishnan, N., Sarabia, J.M. and Mínguez, R. (eds.), *Advances in Mathematical and Statistical Modeling*, Chapter 13, 181-194. ISBN 978-0-8176-4625-7, DOI: 10.1007/978-0-8176-4626-4_14, Birkhauser, Boston.

• Livros

- Pacheco, A.; Tang, L.C.; e Prabhu, N.U. (2008). *Markov-Modulated Processes and Semiregenerative Phenomena*. World Scientific, Singapura, 9 chaps., 224 pp. ISBN 978-981-279-318-8.

• Teses de Mestrado

Título: *Análise de Correspondências Generalizada*

Autor: Sérgio Naito, *naito@sapo.pt*

Orientadora: Margarida Mendes Leal

Título: *Métodos de Captura e Recaptura para a Estimação da Abundância de uma População: Aplicação da Metodologia Bootstrap*

Autora: Catarina Oliveira, *cat_oliveira75@hotmail.com*

Orientadora: Teresa Pereira

Título: *Análise de Regressão e Séries Temporais. Aplicação ao estudo da evolução e previsão do PIB português*

Autor: Rui Santos, ruiavsantos@gmail.com

Orientadora: Teresa Alpuim

Título: *Teoria dos valores extremos, uma abordagem*

Autor: Ivanilda Cabral, ivaneidy@hotmail.com

Orientadora: Sandra Mendonça

Título: *Stochastic processes in Cost-effectiveness analysis in Health Care*

Autor: Marta Ferreira Oliveira Soares, martasoares@hotmail.com

Orientadora: Luísa Canto e Castro Loura

Título: *Modelos Rasch e Escalas de Likert e Thurstone na Medição de Atitudes*

Autora: Luísa Margarida Martins de Campos Antunes da Cunha, lcunha@ispa.pt

Orientadora: Luísa Loura

Título: *Estudo Sobre a Adaptação dos Modelos Log-lineares à Ordinalidade e à Presença de Zeros Amostrais em Tabelas de Contingência*

Autora: Regina Maria Baltazar Bispo Carita, rbispo@ispa.pt

Orientador: Dinis Pestana

Título: *Significância Estatística em Testes de Hipóteses Simultâneos: uma Abordagem Envolvendo Modelos de Mistura*

Autora: Ana Patrícia Subtil da Graça Freitas Garcia, ana_p_subtil@hotmail.com

Orientadora: Luísa Loura

Título: *Modelos de Mutação em Microssatélites e Inferência Filogenética*

Autora: Sandra Isabel Henriques Silva Conceição, sandraihsconceicao@sapo.pt

Orientadora: Luísa Loura

Título: *Estimação da Probabilidade de Ocorrência de Acontecimentos Raros em Contextos Multivariados*

Autora: Inês Filipa Barata Gonçalves, inesfbg@hotmail.com

Orientadora: Luísa Loura

Título: *Modelos de Regressão Dinâmica na Revisão das Séries do Inquérito ao Emprego*

Autora: Ana Luísa de Sousa Lourenço Quitério, anaquiterio@gmail.com

Orientadora: Lucília Carvalho

Título: *Predição de Variantes Genéticas do VIH-1*

Autora: Rute Gomes Velosa Vieira, rvv.gen@gmail.com

Orientadora: Lisete Sousa

Título: *Desenvolvimento de Sistema on-line para a Predição da Topologia de Proteínas Transmembranares: Algoritmo BETH*

Autor: Sérgio Miguel Nunes dos Santos, sergio_santos99@yahoo.com

Orientadora: Lisete Sousa

Título: *Análise Bayesiana de Séries Temporais*

Autora: Ângela Marlene Pires da Cruz, angela.marlene@ua.pt

Orientadora: Isabel Pereira

Título: *Valor Prognóstico da Citometria de Fluxo de ADN no Carcinoma do Colo do Útero*

Autora: Ausenda Pires, 151281@isabelle.math.ist.utl.pt

Orientador: Giovanni Silva

Título: *O Conceito de Profundidade em Estatística*

Autora: Maria Raquel Neto, mariaraquelneto@hotmail.com

Orientador: João Branco

Título: *First Passage Times in Ito Processes Considering a Jump Scenario - Applications in Finance*

Autora: Sílvia Nobre, solarsystem5@sapo.pt

Orientadora: Cláudia Nunes Philippart

Título: *Seleção de Modelos em Amostragem por Distâncias.*

Autora: Maria de Fátima Romão Morgado, romaofatima@sapo.pt

Orientador: Russell Alpizar-Jara

Título: *Validação de Critérios na Avaliação de Qualidade de uma Rede de Transportes Públicos.*

Autor: Sérgio Diogo Gonçalves Santiago, sdgsantiago@gmail.com

Orientadores: Gueorgui Smirnov e Maria do Carmo Miranda Guedes.

• Teses de Doutoramento

Título: *Perdas de clientes e outras medidas de desempenho em sistemas regulares e oscilantes $M^X/G/1/n$ e $GI^X/M(m)//n$*

Autora: Maria Helena Coelho Ribeiro, mhcr@estg.ipleiria.pt

Orientador: António Manuel Pacheco Pires

Na minha tese estudamos filas de espera regulares (i.e., padrão) e oscilantes $M^X/G/1/n$ e $GI^X/M(m)//n$. Nestes sistemas de filas de espera os clientes chegam à fila em grupos segundo um processo geral de renascimento. Os tamanhos dos grupos são variáveis aleatórias independentes e identicamente distribuídas e são independentes do processo de chegadas de grupos de clientes. A capacidade da fila é finita, n , e se à chegada de um grupo de tamanho s só existem r , $r < s$, posições disponíveis na fila, os primeiros r clientes do grupo são aceites e os restantes $s - r$ clientes são rejeitados. Em particular, nas filas $M^X/G/1/n$ a chegada de grupos de clientes ocorre segundo um processo de Poisson e os clientes são servidos por um único servidor. Nas filas $GI^X/M(m)//n$ a taxa de serviço é modulada pelo estado do sistema. Nos sistemas oscilantes $M^X/G/1/n$ e $GI^X/M(m)//n$ o mecanismo de serviço reage à congestão do sistema. Especificamente, os tempos de serviço ou as taxas de serviço oscilam entre duas formas de acordo com a evolução do número de clientes no sistema. Estudamos os sistemas $M^X/G/1/n$ tirando partido da sua estrutura regenerativa markoviana. Usamos cadeias de Markov embebidas e determinamos a distribuição limite do número de clientes no sistema, imediatamente após a terminação de serviço de clientes e em tempo contínuo. Estudamos probabilidades de perdas consecutivas de clientes em períodos de ocupação contínua e momentos da duração de períodos de ocupação contínua. Relacionamos probabilidades de perdas consecutivas em períodos de ocupação contínua e momentos da duração dos períodos de ocupação contínua que se iniciam com múltiplos clientes com as correspondentes quantidades, em períodos de ocupação contínua que se iniciam com um cliente em sistemas com capacidade igual ou inferior. No estudo dos sistemas $GI^X/M(m)//n$ combinamos cadeias de Markov embebidas e uniformização para determinar a distribuição limite do número de clientes no sistema. Calculamos probabilidades de perdas consecutivas em períodos de ocupação contínua que se iniciam com múltiplos clientes como probabilidades de absorção numa cadeia de Markov em tempo discreto. Ao longo da tese propomos algoritmos que implementam os procedimentos apresentados no cálculo das quantidades estudadas e analisamos a sua complexidade.

Helena Ribeiro

Título: *Séries Temporais de Memória Longa com Aplicações ao Controlo Motor – Estudo de Tarefas de Tapping Repetido*

Autora: Ana Maria Fité Alves Diniz, *adiniz@fmh.utl.pt*

Orientadores: Nuno Paulo de Sousa Arrobas Crato e João Manuel Pardal Barreiros

A minha tese baseia-se na caracterização de modelos de séries temporais e na sua aplicação a séries de tarefas motoras. Em particular, a tese incide sobre a modelação de séries de intervalos inter-resposta em tarefas de *tapping*, através de modelos que geram memória longa e que fornecem explicações sobre a presença dessa memória.

Em muitas áreas científicas, um factor importante é a evolução de fenómenos ao longo do tempo. Neste contexto, a abordagem das séries temporais desempenha um papel fundamental, permitindo modelar os fenómenos e fazer inferência. Na área do controlo motor, os estudos de séries relacionadas com a percepção do tempo e com tarefas motoras repetitivas têm revelado resultados típicos de memória longa.

No meu trabalho analisa-se algumas séries temporais de intervalos inter-resposta em tarefas de *tapping* com intervalos-alvo de 800 ms e de 1600 ms. Mais concretamente, efectua-se a análise exploratória das séries, trata-se os *outliers*, estima-se a memória e procede-se à modelação. Neste último passo toma-se como ponto de partida o modelo de Wing-Kristofferson e propõe-se algumas alternativas teóricas em paralelo com as respectivas interpretações biológicas. Por fim refere-se a utilização destas metodologias em séries de ritmos de populações restritas (crianças com hiperactividade e défice de atenção, indivíduos com doenças degenerativas, bebés prematuros) e salienta-se a sua relevância no diagnóstico e na monitorização de doenças.

Ana Diniz

Título: *Tábuas de Mortalidade Contemporâneas e Prospectivas: Modelos Estocásticos, Aplicações Actuarias e Cobertura do Risco de Longevidade*

Autor: Jorge Miguel Ventura Bravo, *jbravo@uevora.pt*

Orientadores: Carlos Braumann e Carlos Pereira da Silva

A minha tese procura constituir-se como uma abordagem integrada e sistemática sobre o problema da medição e gestão dos riscos de mortalidade e de longevidade no contexto actuarial e financeiro e em estudos demográficos. Os significativos ganhos de esperança média de vida registados nas últimas décadas colocam novos desafios em múltiplas áreas, ameaçando em particular a sustentabilidade dos tradicionais sistemas públicos e privados de segurança e protecção social. Nas companhias de seguros do ramo vida, a abordagem tradicional ao cálculo dos prémios e das reservas matemáticas, assente numa intensidade de mortalidade determinística e em taxas de juro constantes, considera-se desajustada nos casos em que a mortalidade é dinâmica. Neste sentido, a dissertação oferece uma revisão detalhada sobre os principais métodos paramétricos e não-paramétricos de graduação de tábuas de mortalidade contemporâneas, incluindo um estudo sobre o seu desempenho na população de pessoas seguras e de beneficiários de fundos de pensões em Portugal. São igualmente analisadas técnicas clássicas de projecção da mortalidade e investigadas novas soluções para a construção de tábuas prospectivas. Os resultados são aplicados na elaboração das primeiras tábuas prospectivas para as populações portuguesa e de pensionistas e na avaliação do efeito selecção adversa. É desenvolvida uma abordagem simultaneamente dinâmica e estocástica do risco de longevidade admitindo que a intensidade de mortalidade pode ser modelada, quer por processos estocásticos do tipo afim, quer por uma função do tipo afim de um conjunto de factores latentes com dinâmica ditada por equações diferenciais estocásticas com saltos. Esta solução permite-nos captar duas características importantes da mortalidade: dependência temporal e incerteza sobre a trajectória futura, e possibilita a derivação de soluções fechadas para a probabilidade de sobrevivência, de fácil aplicação no contexto actuarial. São ainda analisadas um conjunto de estratégias internas e externas de cobertura do risco de longevidade, envolvendo a transferência do risco para os mercados de capitais e a criação de um mercado de derivados de mortalidade.

Jorge Miguel Bravo

Título: *Contributos para o Estudo de Dados em Falta*

Autor: Maria Fernanda Nunes Diamantino, *mfdiamantino@fc.ul.pt*

Orientador: Dinis Duarte Ferreira Pestana

Na minha tese são abordadas questões que têm que ver com o papel importante que as médias ocupam em inferência estatística. Sendo uma questão vasta, a tese foca questões associadas a studentização (com uma incursão de análise de escala em populações não gaussianas) e as questões de não-resposta que se colocam em todas as sondagens.

Foi estudada de uma forma nova a questão da determinação da dimensão da amostra, de forma a torná-la adequada para obtenção da precisão desejada, quando se sabe que muitos dos inquiridos não fornecem a informação solicitada. Neste contexto, atingem-se situações que nos levaram à aplicação da rarefação de Rényi.

Tendo em conta resultados de Rényi sobre rarefação e de Kozubowski sobre somas aleatórias com subordinadora geométrica, investigámos a estimação de parâmetros populacionais usando uma alternativa diferente da solução *ad-hoc* habitualmente explorada: em vez de aumentar a dimensão da amostra deterministicamente para contrariar o efeito de filtragem (não-respostas), estudámos a situação em que, em vez do T.L.C. clássico e aproximação por gaussiana, se pode usar N -divisibilidade infinita no esquema de somas aleatórias subordinadas pela geométrica (geo-somas), e o correspondente limite exponencial ou, mais geralmente, Laplace (no caso de parente com variância finita).

O estudo da qualidade dos estimadores obtidos foi feito usando uma família “dobrada” (*folded*) que generaliza a exponencial e a gaussiana dobrada, associada a uma família “gaussiana generalizada”, já anteriormente usada na investigação de questões de robustez. Foram obtidos números pseudo-aleatórios da família, que denominámos “GLE”, Gaussiana – Laplace Estendida.

Foram ainda estudados alguns aspectos que envolvem os estudantes africanos da FCUL, tendo sido efectuada uma análise do percurso académico destes estudantes feita com base em dados administrativos, e um breve estudo de imputação.

Fernanda Diamantino

Título: *Probabilidade Circa 1914 e a Construção de Pacheco d'Amorim*

Autor: Rui Filipe Vargas de Sousa Santos, *rsantos@estg.ipleiria.pt*

Orientador: Dinis Duarte Ferreira Pestana

Na minha tese é analisada a construção da Teoria da Probabilidade proposta por Diogo Pacheco d'Amorim (1888–1976) na sua tese de doutoramento, intitulada *Elementos de Cálculo das Probabilidades*, defendida em 1914.

A Teoria da Probabilidade é, nesta época, assombrada por diversos paradoxos, originados pela inexistência de uma definição geral de Probabilidade, que ilustram a ambiguidade existente em diversas noções básicas, tais como a escolha *ao acaso*, razões pelas quais David Hilbert (1900) inclui a axiomatização da Probabilidade num dos seus famosos problemas. Pacheco d'Amorim procura responder a este desafio, propondo uma construção rigorosa para a Teoria da Probabilidade baseada no conceito, que considera primitivo, de tiragem *à sorte*, recorrendo a um *fenómeno padrão* onde o agente da selecção procede a escolhas em situação de plena aleatoriedade com total conhecimento do espaço amostra. Deste modo não será necessário recorrer ao polémico *princípio da razão insuficiente* para deduzir a possibilidade de cada elemento. O conceito primitivo de que parte é o de probabilidade condicionada, ainda que a sua definição não seja geral, e concebe o conceito de *ponto imagem* para fundamentar a existência de fenómenos não caracterizados pela equipossibilidade, antecipando muitas ideias subjacentes às variáveis aleatórias e ficando perto de alcançar a definição de função de distribuição. Como epílogo, o autor analisa à luz das leis limites os casos distintos do *fenómeno padrão*, expondo a sua concepção de Estatística.

Neste trabalho são comentadas as principais ideias apresentadas por Pacheco d'Amorim na sua construção, comparando-a com trabalhos da mesma época e outras concepções posteriores, nomeadamente a visão frequencista de probabilidade de Richard von Mises, baseada nos *colectivos*, e a interpretação bayesiana personalista de Bruno de Finetti, centrada no conceito de *permutabilidade*, onde salientamos as principais dificuldades que existiram na procura de uma fundamentação da noção de probabilidade antes da axiomática de Kolmogoroff (1933).

Rui Santos



SOCIEDADE PORTUGUESA
DE ESTATÍSTICA

PRÉMIOS “ESTATÍSTICO JÚNIOR 2009”

Está aberto, até 29 de Maio de 2009, o concurso para atribuição de prémios “**Estatístico Júnior 2009**”, de acordo com o seguinte regulamento:

- 1.** A atribuição de prémios “**Estatístico Júnior 2009**” é promovida pela Sociedade Portuguesa de Estatística (SPE), com o apoio da Porto Editora, e tem como objectivo estimular e desenvolver o interesse dos alunos do ensino básico e secundário pelas áreas da probabilidade e estatística.
- 2.** Os candidatos a prémios “**Estatístico Júnior 2009**” devem ser alunos do 3.º Ciclo do Ensino Básico ou do Ensino Secundário no ano lectivo 2008/2009.
- 3.** As candidaturas podem ser individuais ou em **grupo com um máximo de 3 alunos**. Do grupo pode ainda fazer parte um professor do ensino básico ou secundário ao qual caberá o papel de orientador.
- 4.** Os candidatos devem apresentar um trabalho cuja temática deve estar relacionada com a teoria da probabilidade e/ou estatística.
- 5.** O trabalho deverá ser constituído por um texto escrito em Português com um máximo de 10 páginas A4 dactilografadas e um poster formato A2 que resuma os principais aspectos do trabalho.
- 6.** Poderão ser atribuídos prémios “**Estatístico Júnior 2009**” a 6 trabalhos: aos três primeiros classificados de entre os trabalhos candidatos do 3.º Ciclo do Ensino Básico e aos três primeiros classificados de entre os trabalhos candidatos do Ensino Secundário. Os prémios são constituídos por produtos pedagógicos editados pela Porto Editora (à excepção de manuais escolares) no valor de 600 euros, 300 euros e 200 euros, a atribuir, respectivamente, aos grupos cujos trabalhos sejam classificados em 1.º, 2.º e 3.º lugar.
- 7.** Ao professor orientador do trabalho classificado em 1º lugar, em cada categoria, é ainda atribuída uma anuidade grátis como sócio da SPE, ajudas de custo para participação no XVII Congresso Anual da SPE e produtos pedagógicos editados pela Porto Editora (à excepção de manuais escolares) no valor de 500 Euros.
- 8.** Aos grupos proponentes dos trabalhos classificados em 1º lugar será também oferecida uma ampliação do correspondente poster que será colocado na Sessão de Posters do XVII Congresso Anual da SPE.
- 9.** O boletim de candidatura, acompanhado do trabalho concorrente, deverá ser dirigido ao Presidente da SPE para a morada abaixo indicada. O carimbo do correio validará a data de entrega.

Sociedade Portuguesa de Estatística – Bloco C6, Piso 4 – Campo Grande – 1749-016 Lisboa

O boletim de candidatura e este regulamento podem ser obtidos em

<http://www.spestatistica.pt/static/docs/BoletimCandidaturaPEJ09.pdf>

<http://www.spestatistica.pt/static/docs/RegulamentoPEJ09.pdf>

- 10.** A admissibilidade e apreciação dos trabalhos submetidos a concurso é da competência de um júri, cuja constituição e nomeação será da responsabilidade da Direcção da SPE.
- 11.** O júri é soberano nas decisões, não havendo lugar a impugnação ou recurso.
- 12.** A atribuição dos prémios “**Estatístico Júnior 2009**” será anunciada logo que conhecida a decisão do júri e a sua entrega formal será realizada no XVII Congresso Anual da SPE.
- 13.** Os prémios “**Estatístico Júnior 2009**” poderão não ser atribuídos.

Apoio da Porto Editora



SOCIEDADE PORTUGUESA
DE ESTATÍSTICA

PRÉMIO SPE 2009

Está aberto, até **15 de Junho de 2009**, o concurso para atribuição do **Prémio SPE 2009**, de acordo com o seguinte regulamento:

1. Pretendendo dar destaque ao XVII Congresso Anual da **SPE**, a principal reunião científica organizada pela Sociedade Portuguesa de Estatística, é instituído o **Prémio SPE 2009**.
2. Este prémio destina-se a estimular a actividade de estudo e investigação científica em Probabilidade e Estatística entre os jovens que trabalham nestas áreas.
3. O **Prémio SPE 2009** é constituído por uma quantia de 1000 euros.
4. Ao **Prémio SPE 2009** podem concorrer trabalhos originais sobre temas de Probabilidade e Estatística, desde que não tenham sido objecto de qualquer prémio atribuído por outra instituição.
5. Os candidatos ao **Prémio SPE 2009** devem ser estudantes ou investigadores em alguma instituição portuguesa ou bolseiros portugueses, devem ser sócios da **SPE** e não devem ter completado os 35 anos de idade até 15 de Junho de 2009. Os candidatos não devem ter recebido o Prémio SPE nas quatro edições anteriores.
6. O trabalho deve ser escrito em português e não poderá exceder 25 páginas A4.
7. As candidaturas, o trabalho concorrente e o *curriculum vitae* do candidato, deverão ser dirigidas ao Presidente da **SPE**, em carta registada, para a morada abaixo indicada. O carimbo do correio validará a data de entrega.
8. A admissibilidade e a apreciação dos trabalhos submetidos a concurso é da competência de um júri, cuja constituição será da responsabilidade da Direcção da **SPE**.
9. O júri é soberano nas suas decisões, não havendo lugar a recurso.
10. O trabalho galardoado com o **Prémio SPE 2009** será apresentado pelo seu autor/pela sua autora no XVII Congresso Anual da **SPE** e será publicado nas respectivas Actas.
11. A atribuição do **Prémio SPE 2009** será anunciada logo que conhecida a decisão do júri e a sua entrega formal será feita na Sessão de Encerramento do XVII Congresso Anual da **SPE**.
12. O **Prémio SPE 2009** poderá não ser atribuído.
13. Havendo mais do que um candidato para a atribuição, em igualdade de condições, terá preferência o candidato que apresentar melhor *curriculum* e, subsistindo a igualdade, preferirá o candidato mais jovem.

Sociedade Portuguesa de Estatística
Bloco C6, Piso 4 - Campo Grande
1749-016 LISBOA