

Boletim



SPE

Sociedade Portuguesa
de Estatística

Publicação semestral

outono de 2020



40 anos SPE: De onde viemos? Onde estamos? Para onde vamos?

Probabilidades e Estatística no Departamento de Matemática da Universidade de Coimbra: uma perspetiva	
Esmeralda Gonçalves e Paulo Eduardo Oliveira	14
Alguns tópicos de investigação em estatística realizada no CIDMA – uma breve discussão	
Pedro Macedo, Ana Helena Tavares e Marco Costa	16
Pseudovalorers-p na construção do conhecimento científico	
Maria de Fátima Brillhante	24
Um olhar sobre o desenvolvimento de modelos de sobrevivência para acontecimentos recorrentes	
Ivo Sousa-Ferreira, Ana Maria Abreu e Cristina Rocha	27
O papel da Estatística na Universidade do Porto	
Conceição Nunes Rocha e Isabel Silva	36
Ser SPE – A vantagem do ónus da Boa Estatística	
Jessica Silva Lomba	41
UM departamento com ESTATÍSTICA/Departamento de Matemática/Universidade do Minho (UM)	
Ana Paula Amorim e outros	44
SPE – 40 anos de “braço dado” com a Estatística	
Célia Nunes	48
Estatística na Universidade de Évora	
Russell Alpizar-Jara	51
Uma caminhada por valores extremos de precipitação na Ilha da Madeira	
Délia Gouveia Reis	55
Alguns desafios para o futuro da SPE	
Kamil Feridun Turkman	58
... quarenta anos de sociedade; quarenta e cinco de investigação...	
Tiago A. Marques e Lisete Sousa	60
Testemunhos de Presidentes da SPE	64

Editorial	1
Mensagem da Presidente	2
Notícias	3
<i>Enigmística</i>	7
Episódios na História da Estatística	8
SPE e a Comunidade	10
Ciência Estatística	80

Informação Editorial

Endereço: Sociedade Portuguesa de Estatística.

Campo Grande. Bloco C6. Piso 4.

1749-016 Lisboa. Portugal.

Telefone: +351.217500120

e-mail: spe@spestatistica.pt

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

ISSN: 1646-5903

Depósito Legal: 249102/06

Tiragem: Edição digital

Execução Gráfica e Impressão: Gráfica Sobreireense

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

O papel da Estatística na Universidade do Porto

Conceição Nunes Rocha, *conceicao.n.rocha@inesctec.pt*
CPES (Power and Energy Systems) - INESC TEC, Porto

Isabel Silva, *ims@fe.up.pt*
Faculdade de Engenharia, Universidade do Porto & CIDMA

1 Introdução

Após a surpresa inicial por recebermos o convite para participar neste Boletim SPE, que agradecemos, e olhando para o tema proposto, “40 anos SPE: De onde viemos? Onde estamos? Para onde vamos?”, foi difícil decidir do que falar e por onde começar. Optámos por abordar a investigação em Estatística na Universidade do Porto. Alertamos, desde já, que este trabalho reflete a visão muito particular das suas autoras, que certamente pecará por incompleta e desde já esperamos que qualquer omissão seja perdoada.

Longe vai o tempo em que era suposto os conhecimentos estatísticos dos Engenheiros, dos Psicólogos, dos Químicos, etc., serem suficientes para tratarem os dados e responderem às questões em causa. Na realidade, devido à natureza dos problemas e ao aparecimento de um volume cada vez maior de dados e variáveis, a própria Estatística evoluiu e, como tal, temos mais modelos, mais testes, etc. De uma maneira geral, nas cadeiras de Estatística ou de Análise de Dados lecionadas, os docentes não abordam nem metade dos temas da Estatística (quanto mais dos subtemas), e isso faz com que os estatísticos sejam profissionais essenciais nas equipas que têm por objetivo analisar dados, modelar, prever, etc. No entanto, se antes havia dificuldade em obter amostras que não fossem de reduzida dimensão, hoje para além dessa dificuldade temos também o problema de trabalhar com grandes quantidades de dados. Como, neste último caso, fazer cálculos à mão está fora de questão é reconhecida a importância de um estatístico saber programar, principalmente em R e/ou Python, criando deste modo a necessidade de uma sinergia entre Ciência dos Computadores e Estatística.

Como a nossa colega Conceição Amado (2017) [1] referiu no seu artigo, é cada vez mais importante a

“... necessidade de a Estatística se fortalecer com novas ferramentas teóricas quando a teoria atual se baseia em pressupostos que não são concretizados na prática.”

Adicionalmente,

“O desenvolvimento de metodologias de análise da fiabilidade da informação será fundamental”

quando é feito o

“... tratamento de um enorme volume de dados...”

2 A Estatística na Universidade do Porto

A Universidade do Porto (U.Porto), fundada em 1911, possui 14 faculdades, uma *business school* e mais de 50 centros de investigação.

Embora a U.Porto não forneça nenhum curso específico em Estatística, é na Licenciatura em Matemática¹ ou nos Mestrados em Matemática² ou Engenharia Matemática³ da Faculdade de Ciências da U.Porto (FCUP), que poderá ser iniciado o estudo da Estatística.

Outras alternativas fornecidas ao nível do 2º ciclo é o Mestrado em Modelação, Análise de Dados e Sistemas de Apoio à Decisão⁴ da Faculdade de Economia da U.Porto (FEP), e recentemente o Mestrado em Ciência de Dados⁵ da FCUP (primeira edição em 2018/2019) e o Mestrado em Engenharia e Ciência de Dados⁶ da Faculdade de Engenharia da U.Porto (FEUP), com primeira edição no atual ano letivo 2020/2021.

Ao nível do 3º ciclo, a investigação em Estatística poderá ser continuada, por exemplo, no Doutoramento em Matemática Aplicada (MAP-PDMA, oferecido pelas Universidades do Minho, de Aveiro e do Porto e suportado por 6 unidades de investigação⁷), no Programa Inter-Universitário de Doutoramento em Matemática (da FCUP em conjunto com a Universidade de Coimbra⁸) ou no Programa Doutoral em Ciência de Dados de Saúde⁹ da Faculdade de Medicina da U.Porto.

Parte dos centros de investigação da U.Porto estão integrados nas faculdades ou em laboratórios associados, outra parte foi acolhida em associações de interface em que participam várias entidades públicas e privadas, para além da U.Porto.

Alguns dos centros de investigação da U.Porto são especializados nos mais diversos domínios de conhecimento (desde as Ciências da Vida e da Saúde às Ciências Agrárias, das Engenharias às Belas Artes, não esquecendo as Humanidades) e outros são multidisciplinares. Apresentam dimensões, finalidades e estruturas organizativas muito diversas.

Sediado na FCUP encontra-se o Centro de Matemática da Universidade do Porto (CMUP). O seu Grupo de Probabilidades e Estatística¹⁰, constituído maioritariamente por matemáticos, desenvolve investigação em várias áreas como teoria ergódica, teoria de valores extremos, processamento de sinal e modelagem estatística e inferência. Além disso, são estudadas aplicações em biomedicina, engenharia, seguros e meteorologia.

Na área das Ciências Sociais, podemos mencionar o Centro de Economia e Finanças da Universidade do Porto (cef.up), sediado na FEP, em que uma das áreas de estudo do grupo de investigação THEOMET¹¹ (*Theory & Methods*) é a econometria.

Entre os centros multidisciplinares podemos nomear a Unidade de Investigação em Epidemiologia (EPIUnit), no Instituto de Saúde Pública da Universidade do Porto (ISPUP), cujo Grupo de Bioestatística e Medições em Saúde¹² tem como principal objetivo a investigação em bioestatística, através do desenvolvimento de técnicas matemáticas e estatísticas, com aplicações às áreas de Epidemiologia e Saúde Pública, Medicina e Biologia.

¹<https://matematica.fc.up.pt/licenciatura-em-matematica>

²<https://matematica.fc.up.pt/mestrado-em-matematica>

³<https://matematica.fc.up.pt/mestrado-em-engenharia-matematica>

⁴<https://cursos.fep.up.pt/mestrados/modelacao-analise-de-dados-e-sistemas-de-apoio-a-decisao-mestrado/>

⁵<https://mds.dcc.fc.up.pt/>

⁶<https://dei.fe.up.pt/mdse/>

⁷<https://map-pdma.up.pt/>

⁸http://www.mat.uc.pt/phd_prog/default.php

⁹<https://heads.med.up.pt/>

¹⁰<https://cmup.fc.up.pt/main/content/probability-and-statistics>

¹¹<https://cefup.fep.up.pt/theomet/>

¹²https://isepup.up.pt/research/eniunit/biostatistics_health/

O Centro de Investigação em Tecnologias e Serviços de Saúde (CINTESIS), embora sediado na U.Porto, é uma grande estrutura que inclui 46 instituições parceiras (29 instituições de ensino superior, 12 hospitais ou instituições de saúde e 5 empresas de saúde) e polos em 6 instituições de ensino superior. A terceira linha temática do CINTESIS é denominada Ciência de Dados, de Decisão & Tecnologias de Informação¹³ e é composta por 9 grupos de investigação. Estes grupos são, na sua grande maioria, formados por equipas multidisciplinares de especialistas em áreas como Ciências da Computação, Engenharia, Matemática, Economia e Medicina, cujos objetivos incluem o desenvolvimento de metodologias que permitam contribuir para o processo de tomada de decisão e para a avaliação das intervenções em saúde, assim como para reunir e analisar dados e informação biomédica complexa.

Outra grande instituição de investigação interdisciplinar dedicada à inovação e pesquisa em Ciências da Saúde é o consórcio i3S¹⁴, que resulta de uma longa colaboração entre o IBMC, o INEB e o IPATIMUP. Entre os cursos oferecidos pela Unidade de Formação Avançada do i3S, podemos encontrar um dedicado à Análise Estatística de Dados Biológicos e outro com dois módulos em Estatística Aplicada.

O INESC TEC¹⁵ é um Laboratório de I&D que agrupa investigadores de diferentes Faculdades da U.Porto, do Instituto Politécnico do Porto, da Universidade do Minho, da Universidade de Trás-os-Montes e Alto Douro e de outras universidades internacionais. Na realidade, o INESC TEC é uma instituição de interface que junta a academia, as empresas, a administração pública e a sociedade. Consequentemente, o seu universo é constituído por investigadores de diferentes origens e com diferentes formações. Dizer que a chave do seu sucesso está no facto de se trabalhar em equipa é pobre. Na realidade, as equipas são dinamicamente formadas com base no reconhecimento das diferentes competências individuais e na mais valia de cada um para se atingir os objetivos associados a cada projeto, construindo deste modo equipas multidisciplinares. Como tal, não é de admirar que matemáticos e estatísticos façam parte dessas equipas. No entanto, há um grupo do INESC TEC onde a sua presença é mais forte; estamos a falar do Laboratório de Inteligência Artificial e Apoio à Decisão (LIAAD)¹⁶, pertencendo ao *Cluster* de Ciência da Computação, com uma longa experiência em Inteligência Artificial. A investigação do LIAAD foca-se nas áreas de estatística, otimização, matemática e *machine learning*.

3 Onde estamos e para onde vamos?

Atualmente, na maior parte dos cursos, já se incluiu no currículo da disciplina de Estatística o uso de softwares como o R, o que no nosso entender, resultou do facto de os estatísticos usarem no seu dia a dia essas mesmas ferramentas e não se imaginarem a fazer uma análise estatística sem esse recurso (ou outro análogo). Por outro lado, na U.Porto temos observado o aparecimento de ciclos de estudos que prometem fazer a integração entre as áreas de Estatística e de Ciência de Computadores na formação dos seus estudantes. Mas apesar da expectativa colocada nestes ciclos de estudo para formar os futuros *Data Scientists*, acreditamos que a seleção dos conhecimentos teóricos que devem fazer parte destes ciclos de estudo está incompleta. Uma análise mais cuidada aos programas dos cursos coloca a descoberto o facto de se disponibilizarem muito poucas disciplinas de Estatística em comparação com as de Ciência de Computadores.

Nos últimos tempos, termos como *Data Science* ou *Data Analyst*, entre outros, têm surgido com frequência, o que levou Diggle (2015) [4] a afirmar que:

“The rise of data science could be seen as a potential threat to the long-term status of the statistics discipline. I first argue that, although there is a threat, there is also a much greater opportunity to re-emphasize the universal relevance of statistical method to the interpretation of data.”

¹³<http://cintesis.eu/pt/linhas-de-investigacao/>

¹⁴<https://www.i3s.up.pt/>

¹⁵<https://www.inesctec.pt/pt>

¹⁶

Já Olhede *et al.* (2018) [5] referiu que:

“Much of data science to date has focused on purely predictive “black box” tools rather than classical modeling, inference, and analysis. The role of statistics is to make our understanding of observed phenomena quantitative and precise.”

Nós concordamos com a nossa colega Paula Brito (2016) [3], que considera que:

“Data Science - the science of extracting knowledge from data, não é nem pode ser “apenas” Estatística, mas também não se limita à área da Computer Science. E não pode ser apenas a “colagem” de duas abordagens distintas – requer uma visão integrada, que vai buscar contributos importantes àquelas duas áreas, mas também requer competências de comunicação, de técnicas de visualização, de análise de negócio... Se os estatísticos abraçarem o desafio de se abrirem a novas formas de pensar e de comunicar, a Estatística terá um papel relevante, e que se nos afigura imprescindível, para o desenvolvimento sadio da Data Science.”

E ainda com Regina Bispo (2017) [2]:

“... a análise de dados obriga à dotação de múltiplas competências que vão de facto além das estritamente associadas à formação clássica em Estatística . Assim, tudo (me) indica que estamos perante uma oportunidade de mudar e evoluir, trabalhando em colaboração e complementaridade com outras áreas do saber.”

De facto, o termo Estatística continua a aparecer nos editais publicados para contratar bolseiros ou investigadores, por exemplo no INESC TEC. No entanto, cada vez mais os conhecimentos em Estatística aparecem como fator de preferência e menos como formação base. É comum encontrar propostas de trabalho em que se pede que o candidato tenha:

- formação em ciência dos computadores (ou área afim) e bons conhecimentos de estatística,

ou

- formação em matemática ou estatística (ou área afim) e que saiba programar bem em R ou Python.

Estas opções não são equivalentes, nem tão pouco é indiferente qual destas opções é que faz parte do edital, mas isto é um reflexo do pensamento/necessidade da sociedade que a Estatística tem de acompanhar.

Resumindo, na U.Porto temos estatísticos e matemáticos a trabalhar em colaboração com colegas de diversas áreas e em diversas aplicações, fortalecendo deste modo o trabalho científico das equipas. No entanto, acreditamos que o futuro da Estatística na U.Porto, e não só, não se pode limitar a dar formação estatística a graduados em ciência dos computadores ou vice-versa, sendo importante criar um perfil novo. A sociedade necessita de cientistas dos dados, ou seja, formandos com um perfil que preencha todos os requisitos explícitos e implícitos atrás referidos. No nosso entender, o futuro da Estatística na U.Porto passa por criar, de raiz, cursos que formem estes novos perfis científicos.

Assim, ainda há muito a fazer. Voltando ao tema do Boletim e como a nossa colega Conceição Amado (2017) [1] tão bem referiu no seu artigo:

“Às Sociedades de Estatística será requerido um esforço adicional na direção da promoção de interações no “ecossistema” da Estatística, da Computação (hardware e software), da Análise Numérica e da Otimização.”

Um bem haja à SPE!

Referências

- [1] Amado, C. (2017). Espelho meu, espelho meu, que futuro terei eu? *Boletim SPE outono de 2017*, **2**, 32–34.
- [2] Bispo, R. (2017). A tirania dos jargões. *Boletim SPE outono de 2017*, **2**, 38–41.
- [3] Brito, P. (2016). Data Science – um desafio para os estatísticos? *Boletim SPE outono de 2016*, **2**, 40-42.
- [4] Diggle, P.J. (2015) Statistics: a data science for the 21st century. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **178.4**, 793–813.
- [5] Olhede, S.C. e Wolfe, P.J. (2018). The future of statistics and data science. *Statistics & Probability Letters*, **136**, 46–50.

