

Reflexões estatísticas

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

CIMA - Centro de Investigação em Matemática e Aplicações, Instituto de Investigação e Formação Avançada, Universidade de Évora

Agradecendo o convite do nosso amigo Fernando Rosado para escrever para o Boletim da SPE, venho, em jeito de comemoração do merecido sucesso do Boletim e de preocupação com o futuro da SPE e da Estatística, partilhar convosco a minha reflexão sobre três temas. O primeiro está relacionado com o ensino da Estatística. O segundo, pelo qual farei uma passagem fugaz, refere-se a alguns dos desafios atuais da Ciência Estatística que mais têm captado a minha atenção. O terceiro é sobre a comunidade estatística nacional, a sua inserção internacional e o seu futuro.

As recentes polémicas sobre o uso devido ou indevido do valor- p (“ p -value” em inglês) chamam a atenção para os cuidados que há a ter na transmissão aos nossos estudantes das ideias estatísticas. Os riscos de uma abordagem demasiado apressada e superficial são porventura ainda maiores nas unidades curriculares destinadas à formação estatística básica de estudantes de outras áreas. Essa formação é extremamente importante para a utilização e compreensão das análises estatísticas que irão encontrar na sua vida profissional, pelo que, apesar de termos de meter o Rossio na Rua da Betesga, é importante dedicar algum tempo à correta interpretação dos conceitos básicos da Estatística, até porque a sua utilização em unidades curriculares posteriores de outras áreas vem frequentemente reforçar os erros de perceção inicial mais comuns. A situação agrava-se quando se criam unidades curriculares com variadas designações (“Metodologias de Investigação em ...” é muito comum) que não são mais do que unidades curriculares de Estatística assim transvertidas para justificar serem lecionadas (sem qualquer colaboração de professores da área da Estatística) por colegas professores de outras áreas sem formação estatística adequada. Não falando já do problema da interpretação menos correta dos resultados estatísticos, podia dar (mas não vou fazê-lo) exemplos chocantes de utilizações grosseiras de metodologias estatísticas totalmente inadequadas ao problema em estudo, utilizações que obviamente o *software* permite e que, ainda por cima, são às vezes transmitidas às vítimas inocentes (os alunos) como se de grande inovação científica se tratasse.

Mas a reflexão sobre o uso dos valores- p leva-me a uma outra questão que tem a ver com a própria teoria estatística que transmitimos, designadamente com a hipótese nula e os testes de significância. De facto, aceitar ou rejeitar a hipótese nula tem pouco a ver com ela ser verdadeira ou falsa. Não estou a falar de erros de primeira ou segunda espécie, nem sequer dos erros de terceira espécie (inadequação do modelo estatístico). Estou a falar de uma falácia fundamental, podíamos chamar-lhe um erro de quarta espécie, que está subjacente a esta teoria estatística. Com efeito, na esmagadora maioria (senão mesmo na totalidade) das aplicações, mesmo sem quaisquer dados, e obviamente sem qualquer teste estatístico, é seguro afirmar categoricamente e *a priori* que a hipótese nula é falsa. Por que razão é que dois tratamentos de uma doença teriam exatamente a mesma probabilidade de cura, sem o mínimo desvio? Por que razão é que dois materiais diferentes teriam exatamente a mesma resistência? Por que razão é que a altura média μ_1 de uma população de um país haveria de ser exatamente igual à altura média μ_2 da população de outro país? Por que razão dois métodos de ensino deveriam ter exatamente a mesma taxa de sucesso? Mesmo em casos aparentemente mais controlados, por que razão é que a probabilidade de sair cara no lançamento de uma moeda deveria ser exatamente 1/2, isto é, será que há mesmo moedas absolutamente honestas? É óbvio que haverá sempre uma diferença, por mínima que seja. Presumir que seria possível uma igualdade perfeita (entre duas quantidades ou entre uma quantidade e um valor de referência) constitui um modelo teórico que peca por não ser genérico e que

seria liminarmente descartado *a priori* noutras áreas de aplicação de modelos matemáticos, como sucede por exemplo nos sistemas dinâmicos aplicados.

Assim, na esmagadora maioria dos casos, quando testamos uma hipótese nula, concluir pela sua aceitação ou rejeição nada tem a ver com ela ser verdadeira ou falsa. Ela é, como vimos, falsa. Aceitar ou rejeitar a hipótese nula tem apenas a ver com o tamanho da amostra e o acaso. Se a nossa amostra for pequena, a probabilidade de aceitar a hipótese nula é muito elevada, não obstante essa hipótese ser falsa. Se a amostra for grande, a probabilidade de rejeitar a hipótese nula é muito elevada (de facto, esta probabilidade tende para um quando o tamanho da amostra tende para infinito). É claro que a hipótese nula, apesar de falsa, pode ser um instrumento útil, se admitirmos que não exigimos uma igualdade perfeita, isto é, se, para efeitos práticos, considerarmos verdadeira a hipótese nula caso a diferença entre as duas quantidades em causa seja muito pequena (inferior a uma certa margem de tolerância $\varepsilon > 0$). Mas isso não é normalmente dito aos alunos. Designando por d o parâmetro que representa a diferença entre os parâmetros em comparação (por exemplo, $d = \mu_2 - \mu_1$ ou $d = p - 1/2$), o que nós de facto queremos não é testar $H_0: d = 0$ versus $H_1: d \neq 0$ mas decidir entre $H_0: |d| < \varepsilon$ e $H_1: |d| \geq \varepsilon$. O leitor fará as necessárias adaptações se o teste for unilateral em vez de bilateral.

O caso da saúde é paradigmático. Se testarmos um novo medicamento comparativamente ao medicamento anteriormente usado numa determinada patologia, sendo p_1 e p_2 as taxas de cura do antigo e do novo medicamento, respetivamente, testamos $H_0: p_1 = p_2$ (ou seja $H_0: d = p_2 - p_1 = 0$) versus $H_1: p_2 > p_1$ (ou seja $H_1: d > 0$), apesar de H_0 ser falsa à partida. Se o novo medicamento for ligeiramente melhor (isto é, p_2 for maior que p_1 mas por muito pouco), é só fazer ensaios clínicos com amostras suficientemente grandes para concluir que é de rejeitar a hipótese nula e de aceitar H_1 . Se a análise ficar por aí, podemos estar a pressionar as autoridades de saúde para adotarem o novo medicamento, ainda que o seu custo seja exorbitante em comparação com o antigo, quando a sua taxa de cura é apenas muito ligeiramente superior. O que talvez devêssemos fazer era definir uma margem $\varepsilon > 0$ que levaria à adoção do novo medicamento se houvesse evidência de que a sua taxa de cura superava a do antigo no mínimo por essa margem. Deveríamos então decidir entre $H_0: d < \varepsilon$ e $H_1: d \geq \varepsilon$. E note-se que aqui a margem ε pode ser muito pequena mas também pode não ser tão pequena assim, dependendo a decisão política de considerações sobre a eficácia da utilização dos recursos e sobre a gravidade da patologia.

Esta área da saúde é muito delicada mas considerações semelhantes surgem noutras áreas em que a decisão sobre o valor de ε pode ser bem mais simples.

Nada disto é propriamente novidade, antes constitui conhecimento comum da comunidade estatística. Mas raramente passamos este conhecimento aos nossos alunos, até porque o problema reformulado é mais complexo. Mas devíamos, pelo menos, informar que a hipótese nula é uma base de trabalho um tanto “artificial”, sendo importante, além, da decisão sobre a sua aceitação ou rejeição, fazer intervir nas decisões de atuação outras considerações como, por exemplo, no caso dos medicamentos, as estimativas das taxas de cura.

Sobre o segundo tópico, os desafios atuais da Ciência Estatística, muito haveria a dizer, mas limitar-me-ei a uma breve referência a duas áreas que me parecem particularmente interessantes. Uma é a das grandes bases de dados (“big data sets”) que hoje abundam e do “data mining” (mineração de dados), onde estamos em competição com a informática e a inteligência artificial e onde uma síntese útil entre as abordagens estatística e informática parece ser um desafio interessante. A outra é a da estatística de processos estocásticos (estimação paramétrica, não-paramétrica ou semiparamétrica, simulação de Monte Carlo, seleção de modelos, testes de hipóteses), onde tem havido progressos muito relevantes mas onde há problemas particularmente interessantes que, pela sua dificuldade (especialmente no caso de processos em tempo contínuo em que as observações decorrem em tempo discreto ou quando há variáveis não observadas), ainda têm soluções pouco satisfatórias. Um problema particularmente difícil, mas que seria muito útil desenvolver, tem a ver com os tempos de passagem dos processos por certos limiares, seja o problema de determinar a sua distribuição e/ou os seus momentos, seja o de utilizar a distribuição ou os momentos empíricos dos tempos de passagem para inferir sobre o processo estocástico subjacente.

Finalmente, queria fazer uma referência à evolução da comunidade estatística nacional e aos desafios futuros que enfrenta. Hesitei muito em falar sobre este tema pois ele interliga-se com a atividade da SPE e entendo que quem, como eu, foi Presidente da SPE, deve exercer alguma cautela nas suas

intervenções, não vão elas ser interpretadas como pressão indevida sobre os atuais órgãos sociais. Mas, para além de não querer frustrar o convite do Editor do Boletim, sei que os detentores dos atuais órgãos me conhecem e sabem que eu conheço a sua independência, pelo que não há o risco de interpretarem mal a minha intervenção. Peço desde já desculpa se ela é, como não podia deixar de ser, influenciada pelo envolvimento afetivo de um sócio fundador que teve responsabilidades diretivas e deseja o melhor para a SPE.

A comunidade estatística portuguesa, trabalhando na academia e nos setores público e privado, teve um extraordinário desenvolvimento nas últimas décadas, acompanhando a crescente necessidade de pessoas com formação estatística qualificada. Isso resultou num aumento do número de estudantes a procurar formação inicial e avançada nestas áreas, o que, juntamente com o reconhecimento da necessidade de alguma formação estatística nos mais diversos cursos, levou ao crescimento do setor da Estatística nas instituições de ensino superior, acompanhado de um incremento exponencial da investigação e do reforço quantitativo e qualitativo dos centros de investigação com atividade nesta área. E com isso, e também pela sua elevada qualidade, a comunidade estatística nacional tem-se afirmado, quer no País, onde desempenha um importante papel no desenvolvimento, quer a nível internacional, seja em termos de publicações, seja em termos de reconhecimento, com estatísticos portugueses ocupando por mérito próprio cargos diretivos em sociedades científicas internacionais.

Mas este progresso da comunidade estatística nacional é indissociável da atuação da SPE, que tem tido um papel fundamental na construção do próprio sentido de comunidade e da sua coesão e como catalisador do desenvolvimento da Estatística a nível científico e profissional e na sua internacionalização. O mérito da SPE deve-se naturalmente ao dinamismo e empenhamento dos seus sócios na vida e na atividade da Sociedade.

A SPE tem promovido a cooperação com outros organismos com relevante atividade estatística, como o INE e o Banco de Portugal, e a interação entre as instituições de ensino superior e as instituições públicas e privadas onde a estatística é um instrumento importante de análise e intervenção. A SPE estabeleceu ligações com outras sociedades científicas congéneres, quer a nível europeu (que incluiu, entre outras iniciativas, um acordo para publicação internacional das Atas dos congressos e a criação da FENStatS), quer com o Brasil, quer ainda no âmbito do ISI, que contribuíram para uma maior conetividade internacional. A SPE, em cooperação com outras entidades, tem promovido a divulgação da Estatística, particularmente junto dos jovens, com atividades variadas, como os Prémios Estatístico Júnior, a Radical Estatística ou a Explorística, que é uma exposição interativa itinerante resultante da cooperação da SPE com o Ciência Viva que mereceu já dois prémios internacionais de grande relevo (além de estar a ser exportada para outros países).

Mas a atividade da SPE que mais se destaca, pelo papel fundamental que desempenhou na coesão da comunidade estatística nacional e no desenvolvimento da sua atividade estatística, foi a realização dos Congressos anuais e publicações associadas (atas e livros de minicurso). Os Congressos, para além de constituírem uma vivida e desejada celebração da comunidade e do seu labor já consagrada nos seus hábitos, permitiram um melhor conhecimento mútuo e um incentivo ao desenvolvimento científico e aperfeiçoamento profissional, inclusivamente pelos estudantes de pós-graduação. Claro que a isso não é alheio o convívio num espírito de cordialidade e amizade que as atividades da SPE e a efetiva participação dos sócios na definição e construção do futuro da Sociedade proporcionam.

Recentemente, o País foi abalado por uma crise da qual a Estatística nacional não saiu incólume. A procura da formação avançada em Estatística diminuiu, as instituições de ensino superior passaram de um ciclo de expansão para um de contração, praticamente parando o influxo de sangue novo, o financiamento da investigação sofreu cortes brutais, a investigação estatística e vários dos centros onde decorria foram vítimas de uma política de avaliação incompetente em que a Estatística foi uma das áreas discriminadas (a nível da constituição dos painéis de avaliação e sua valorização do trabalho científico estatístico e a nível do estabelecimento de quotas cegas que não tiveram em conta o grande desenvolvimento desta área). Naturalmente, a SPE sofreu também os efeitos induzidos por esta crise. Em minha opinião, a Sociedade deveria ter tentado, em contraciclo, contrariar esses efeitos, tirando partido da dinâmica positiva que sempre a caracterizou, em vez de tentar acomodá-los aumentando o intervalo entre Congressos (que deixaram de ser anuais, o que foi parcialmente compensado com a realização de outro evento científico intercalar de carácter mais setorial) e eliminando a publicação internacional das Atas. Esta acomodação afetou, na minha opinião, a atratividade dos Congressos e a coesão da comunidade estatística, efeitos que correm o risco de se autoalimentarem. É certo que se

temia o risco de retração das receitas da SPE por efeito da crise, mas o elevado capital financeiro acumulado no passado constituía uma almofada que permitia e permite perfeitamente um investimento no futuro. O vínculo da comunidade estatística nacional à SPE, que se reflete no próprio sentido de comunidade, é hoje provavelmente menos forte que no passado. Precisamos de contrariar esta tendência.

Felizmente, parece-me que as circunstâncias externas que afetaram a comunidade estatística estão lentamente a começar a inverter-se. O setor produtivo dá alguns sinais de retoma, a importância de profissionais com formação estatística qualificada nas organizações públicas e privadas é crescentemente reconhecida, a retoma da procura dessa formação, ainda que com um fator de atraso, começa a dar sinais positivos a que as instituições de ensino superior começam timidamente a responder, a política de avaliação começa a ter outra atitude e a reconhecer a importância da investigação estatística e dos centros de investigação em que se realiza. São ainda sinais débeis, mas que alimentam a esperança.

O vínculo dos sócios à SPE terá sido afetado mas, na maioria dos casos e como se pôde constatar no Congresso de Olhão, é grande a vontade dos sócios em o retomar na plenitude. Se as circunstâncias internas da SPE derem uma ajuda, contribuindo, através de políticas adequadas, para revigorar este vínculo (processo que não é fácil e requer atuação imediata e vários anos de esforços continuados), a comunidade estatística será ainda mais forte e coesa. É esse o desafio da nova Direção e demais órgãos da SPE, bem como de todos os sócios.

Termino formulando votos de continuado sucesso para o Boletim da SPE, para a SPE e para os seus sócios, ao serviço da Estatística e do seu relevante papel no desenvolvimento económico e social.