Estatísticos que Mudaram a Estatística

Dinis Pestana
CEAUL — Centro de Estatística e Aplicações da Universidade de Lisboa
CFCUL — Centro de Filosofia das Ciências da Universidade de Lisboa

FCT Este trabalho é financiado por Fundos Nacionais através da FCT – Fundação para a Ciência e a Tecnologia no âmbito do projecto PEst-OE/MAT/UI0006/2011

Nota prévia: estas resenhas biográficas, e focando a atenção apenas na componente estatística da multifacetada actividade destes génios, é naturalmente muito resumida e simples, pois resulta de uma "encomenda" da professora Doutora Emília Athayde, da Universidade do Minho, que teve a interessante ideia de publicar, com o apoio do INE, um "calendário" para celebrar 2013 — Ano Internacional da Esatística. Escolheu, um ou dois notáveis da Probabilidade e da Estatística nascidos em cada um dos meses do ano. Esperamos do êxito desta iniciativa uma continuação que inclua muitos outros estatísticos que mudaram os rumos da Estatística e a vastidão das suas aplicações, e também o seu papel incontornável na formação da metodologia da investigação em ciências experimentais, sociais e humanas.

Jacques Bernoulli

que também usou os nomes Jacob e James, nasceu em 1654 e morreu em 1705, em Basileia, para onde a família paterna tinha fugido, dos Países Baixos, para escapar à perseguição, ordenada pelo rei de Espanha em 1567, movida contra os que se opunham ao domínio espanhol ou seguiam a fé protestante.

Por imposição paterna, Jacques Bernoulli fez estudos de Teologia. Mas a sua paixão era a Matemática. Uma estadia de dois anos em França permitiu-lhe estudar com discípulos de Descartes. Posteriormente esteve na Holanda e em Inglaterra, conhecendo muitos cientistas interessantes do seu tempo, com quem manteve correspondência científica — como era então hábito comunicar o progresso científico — ao longo da vida.

Depois do seu regresso à Suissa foi ensinar Mecânica na Universidade de Basileia, sendo nomeado professor de Matemática em 1687, posto que ocupou até ao fim da vida. Ensinou Matemática ao seu irmão mais jovem Johann I (1667–1748), que o pai tinha destinado a uma carreira médica, e durante os anos em que colaboraram clarificaram as ideias de Leibniz sobre cálculo infinitesimal, que este expusera de forma muito obscura para a maioria dos matemáticos seus contemporâneos. A colaboração entre os dois irmãos transformou-se porém numa rivalidade e animosidade com troca de "mimos"públicos que não abonam a favor do carácter de nenhum deles[(1)].

A primeira publicação de Jacob Bernoulli sobre Probabilidade data de 1685, e em 1689 revolucionou a probabilidade e as suas aplicações com um teorema a que se referia com ternura como o seu *teorema de ouro*, que actualmente é conhecido como Lei Fraca dos Grandes Números[(2)]. Este primeiro teorema limite pode ser explicado de forma não rigorosa mas simples e apelativa notando que se considerarmos a sucessão de médias $\overline{x}_n = \frac{1}{n}\sum_{k=1}^{n} x_k$ de amostras de tamanho crescente esta sucessão estabiliza porque $\overline{x}_{n+1} = \frac{1}{n+1}\sum_{k=1}^{n+1} x_k = \frac{n}{n+1}\overline{x}_n + \frac{x_{n+1}}{n+1} \approx \overline{x}_n$, assumindo-se que aproximando-se da média populacional μ.

Em termos mais rigorosos e para sucessões de variáveis aleatórias independentes e identicamente distribuídas, há evidentemente que assumir a existência de valores médios $\mathbb{E}[X_k]$, o que garante que $\mathbb{P}[\frac{X_{n+1}}{n+1} > \varepsilon] \to 0$. Ora se considerarmos uma sucessão de variáveis aleatórias independentes "de Bernoulli", $X_k = \left\{ \begin{array}{cc} 0 & 1 \\ 1-p & p \end{array} \right.$, interpretando 0 e 1 como não se ter realizado ou ter-se realizado o acontecimento A, respectivamente, estamos na situação excepcional $\mathbb{E}[X_k] = p = \mathbb{P}[A]$, e consequentemente o teorema de ouro de Bernoulli justifica uma interpretação frequencista da probabilidade, como limite da sucessão de frequências relativas do resultado A quando se realizam inúmeras *provas de Bernoulli*, isto é experiências aleatórias independentes dicotómicas (em que só interessa se se observa ou não A), admitindo que a probabilidade de A se mantém um valor p constante em todas as experiências.

Antes desta interpretação da probabilidade, as aplicações deste ramo da Matemática estavam confinadas a situações de perfeita simetria do universo, permitindo admitir equiprobabilidade dos acontecimentos elementares — essencialmente jogos de azar, com um equilíbrio artificiosamente manufacturado, em que se admitia que o uso de dados não viciados, ou a distribuição honesta, ao acaso, de mãos de cartas de jogar. A nova interpretação frequencista tornou possível aplicar a Probabilidade a estudo de fenómenos naturais em que não era necessário admitir equiprobabilidade, havia apenas que, ao menos conceptualmente, assumir a frequência relativa numa amostra de grande dimensão como uma boa aproximação da probabilidade. É isto que justifica, por exemplo, o notável trabalho (1693) de Halley sobre anuidades, um antepassado das tabelas de mortalidade e de toda a ciência actuarial, que permitiu ao governo britânico vender seguros de vida fixando os preços objectivamente justos com base nas probabilidades de sobrevivência do comprador. A nova interpretação frequencista da probabilidade está também na base do que é actualmente considerado o primeiro teste de hipóteses, uma famosa carta de John Arbutnot (1710) à Royal Society, argumentando a existência da divina providência com base na evidência de que é mais provável que uma mulher dê à luz um rapaz do que uma rapariga.

Note-se que evidentemente Bernoulli não descartou as situações em que se pode

(1) Johan devia ser particularmente insuportável e mesquinho; quando a Academia de Paris partilhou a medalha de Matemática entre ele e o seu filho Daniel, expulsou-o de casa e nunca mais lhe falou! Em 1739 publicou uma *Hydraulica* plagiando uma *Hydrodinamica* deste seu filho Daniel, e inclusivamente antedatando-a de 1732 para fingir prioridade, mas o plágio foi publicamente exposto.

(2) Poisson, ao descobrir que uma sucessão de variáveis aleatórias binomiais com probabilidade $p_n \to 0$ mas com valor médio estabilizado ($n\,p_n \to \mu > 0$), tentou chamar a este resutado "lei dos pequenos números", referindo-se então ao teorema de Bernoulli como "lei dos grandes números".

admitir que os acontecimentos de base são equiprováveis, sendo originalmente dele o *princípio da razão insuficiente* muitas vezes atribuído a Laplace: se não houver razão suficiente para duvidar da equiprobabilidade dos acontecimentos elementares, é razoável e pragmático usar essa hipótese de equiprobabilidade.

As descobertas de Jacques Bernoulli na área de Probabilidade vieram a ser compilados e publicados postumamente, em 1713, pelo seu sobrinho Nikolaus Bernoulli, também ele um probabilista que estabeleceu resultados interessantes. Julga-se que o título *Ars Conjectandi* é uma homenagem à famosa *L'art de Penser*, a famosa *Lógica de Port-Royal* assinada por Arnault e Nicole, mas que se pensa ter tido uma importante contribuição de Pascal.

A primeira parte desta obra é a republicação comentada do folheto de Christiaan Huyghens *De ratiociniis in ludo aleae* (1667), considerado o primeiro livro de Probabilidade. Na segunda parte estão registados muitos resultados importantes sobre permutações e combinações, e é também estudada uma sucessão de números, actualmente denominados *números de Bernoulli*, que tem importantes aplicações em muitos ramos de Matemática. A terceira parte versa aplicação da Probabilidade na avaliação dos riscos das apostas em jogos (em geral incompletamente definidos). É na quarta parte que os pensamentos mais originais e inovadores de Jacques Bernoulli são apresentados — nomeadamente a Lei dos Grandes Números e as suas consequências na interpretação frequencista do conceito de Probabilidade. Não só é um marco na Probabilidade, como na didáctica da Matemática, pela clareza e elegância da exposição dos temas que aborda.

No que respeita à Lei dos Grandes Números, dois comentários parecem inevitávies:

1. A sucesssão de valores médios aproxima-se do valor médio populacional, mas a que ponto a aproximação é boa? — a resposta mais geral é o Teorema limite Central, que mostra que a velocidade de convergência é da ordem de $\frac{1}{\sqrt{n}}$. Resultados mais modernos usando técnicas sofisticadas (aproximações em selas, por exemplo) mostram que esta velocidade de convergência pode ser melhorada para a ordem de $\frac{1}{n}$ em muitas questões em que é necessário ter boas aproximações sem gastar muito tempo (e dinheiro) na repetição de muitas experiências.

2. Não é ainda suficientemente reconhecido que o fundamento teórico de um método de cálculo rápido proporcionado pelo advento de computadores, que Ulam baptizou com o nome de *Método de Monte Carlo*, é uma simples consequência da Lei dos Grandes Números.

Abraham de Moivre

Abraham Moivre (1667-05-26 a 1754-11-27) nasceu em França, numa família protestante, num período de perseguições religiosas que levaram à extinção da Academia Protestante de Sedan em que começou por estudar, o que o obrigou a inscrever-se para estudar Lógica em Saumur, uma instituição católica. Apesar de a Matemática não ser uma das áreas curriculares, por gosto e iniciativa própria leu várias obras dessa disciplina, incluindo *De ratiociniis in ludo aleae*Êde Huygens, considerado o primeiro livro dedicado à Probabilidade.

A família foi viver para Paris, e os estudos de Abraham Moivre continuaram no Collège de Harcourt, onde pela primeira vez teve instrução formal em Matemática. Mas além disso teve lições particulares do famoso pedagogo Ozanan, o que provavelmente viria a influenciar o seu modo de vida, e mesmo o seu interesse por Probabilidade, pois Ozanan adorava jogos de azar, perdendo e ganhando muito dinheiro em apostas.

A Revogação do Edito de Nantes fez recrusceder, e agora com apoio legal, a perseguição dos protestantes em França, e Abraham Moivre chegou a estar encarcerado, não se sabe durante quanto tempo. Essa perseguição originou uma diáspora de protestante para outros países, e Abraham Moivre foi viver para Inglaterra. O seu modo de vida foi dar lições particulares, deslocando-se a casa dos alunos — foi provavelmente para impressionar potenciais clientes que passou então a intitular-se Abraham de Moivre) — ou em "casas de café". Tendo folheado em casa do Conde de Devonshire os *Principia* de Newton, percebeu que esta obra ia muito além dos seus conhecimentos. Conta-se que adquiriu a obra, e a separou em maços de folhas, o que lhe permitia estudá-la nos intervalos entre as explicações que eram o seu modo de vida.

Em 1692 tornou-se amigo de Halley, e posteriormente do próprio Newton. Foi Halley quem apadrinhou a publicação do primeiro trabalho de A. de Moivre, em 1695, sobre as fluxões estudadas nos *Principia* de Newton, nas *Philosophical Transactions of the Royal Society.* A sua generalização do binómio de Newton para uma situação multinomial mereceu amplo reconhecimento, e levou a que se tornasse membro da *Royal Society* em 1697. A "fórmula de de Moivre"para $(cos\, x + \mathrm{i}\, sen\, x)^n$ (que Euler viria a demonstrar de forma mais simples em 1749) estabeleceu a importância de considerar a interligação de números complexos e de trigonometria.

A *Royal Society* tinha encomendado a Halley um relatório sobre anuidades, o que levou a publicar a primeira tabela de mortalidade em 1693. A. de Moivre generalizou de forma não trivial os resultados, usando uma aproximação por funções lineares em vez de usar directamente a tabela de mortalidade de Halley, e curiosamente refere como uma das bases do seu trabalho, no livro *Annuities upon Lives* de 1725 a "probabilidade da duração da vida". Halley e de Moivre são muito justamente considerados os pais do cálculo actuarial.

Fez parte da comissão da Royal Society encarregada de se pronunciar sobre a prioridade de Newton ou de Leibniz no que se refere ao cálculo infinitesimal, o que porventura o prejudicou na obtenção de um contrato universitário no continente; e por outro lado, em 1711 Newton recomendou Nicholas Saunderson para a prestigiosa cátedra Lucasiana de Cambridge, a que de Moivre também se tinha habilitado, o que certamente o fez perceber que tinha que se resignar à pobreza, sobrevivendo com a remuneração das suas lições particulares.

Uma polémica com o escocês Cheyne sobre prioridade de descobertas sobre as fluxões de Newton levou-o a corresponder-se com Johann Bernoulli, percebendo que não dominava os novos métodos de cálculo infinitesimal como este. Foi provavelmente isso, e o interesse dos seus alunos por jogos de azar, que o levou a investir no cálculo de Probabilidades, um assunto novo em que havia a oportunidade de se distinguir. Em 1711 publicou nas *Philosophical Transactions of the Royal Society* com uma primeira versão, *De Mensura Sortis*, em latim, do que viria a ser a sua obra magistral, *The Doctrine of Chances*, que teria em vida do autor edições aumentadas, em 1718, em 1738 e em 1756. A sua convicção do entrosamento indisociável de acaso e necessidade, lapidarmente expresso na abertura de *The Doctrine of Chances*, de 1718

> "*Further, The same Arguments which explode the notion of Luck may, on the other side, be useful in some Cases to establish a due comparison between Chance and Design: We may imagine Chance and Design to be as if it were in Competition with each other, for the production of some sorts of Events, and may calculate what Probability there is, that those Events should be rather owing to one than to the other.*"

é decerto partilhada por muitos cientistas e pensadores da Ciência.

Foi na edição de 1738 que apareceu a primeira forma (muito limitada, uma vez que apenas se estabelece o resultado a possibilidade de aproximação de probabilidades binomiais usando um integral que tem como função integranda o que actualmente consideramos a função densidade de probabilidade da lei normal) do Teorema Limite Central. A. de Moivre considerava, muito justamente, que era um progresso sobre a Lei dos Grandes Números de Jacques Bernoulli (e de facto, em certo sentido e em contextos muito mais gerais, o Teorema Limite Central pode ser perspectivado como uma avaliação simples da velocidade de convergência para o valor médio, que a Lei dos Grandes Números estabelece). A edição de 1756 contém também a segunda parte de *Annuities on Lives*, continuação do seu livro de 1725.

O investimento de A. de Moivre na Probabilidade não esteve liberto de dissabores e polémicas sobre prioridades, primeiro com Montmort e com Nikolaus Bernoulli, mais tarde com Thomas Simpson, que publicou em 1740 e em 1742 livros que competiam com obras de A. de Moivre. No que se refere a Montmort, que em 1708 publicara uma primeira edição do *Essay d'Analyse sur les Jeux du Hazard*, é justo comentar que esta apresentava a solução de muitos problemas que continuam a ser exercícios de modernos manuais de Probabilidade; no entanto nesta primeira edição, na generalidade dos casos, não se apresenta uma base teórica em que se baseava a estratégia usada para resolver o problema. É provavelmente por isso que Montmort é muitas vezes secundarizado nas monografias de história da Probabilidade. Na segunda edição, de 1714, provavelmente como reacção à publicação de A. de Moivre de 1711, há já um enquadramento teórico adequado, e Montmort merece figurar entre um dos pioneiros da Probabilidade. Mas as estrelas desse período inicial são de facto Jacques Bernoulli e Abraham de Moivre, cuja *The Doctrine of Chances* é um marco que se destaca até à publicação em 1812 do *Traité Analytique des Probabilités* de Laplace, e do seu famoso prefácio *Essai Philosophique sur les Probabilités* à reedição de 1814 (e que também teve, nesse mesmo ano, edição separada).

Tal como Cardano, o autor do primeiro livro tratando matematicamente jogos de azar, A. de Moivre previu exactamente a data da sua morte. Diz-se que tendo observado no fim da vida que estava a dormir em cada dia mais 15 minutos, e que assim previu que em 27 de Novembro de 1754 já não acordaria.

Laplace

Pierre-Simon, Marquês de Laplace (1749-03-23 a 1827-03-05), muitas vezes apodado de “o Newton francês”, porventura devido à sua monumental *Tratado de Mecânica Celeste*, precedido de uma introdução não matemática *Exposição do Sistema do Mundo*, distinguiu-se em muitas áreas da Matemática, e nomeadamente revolucionou a Probabilidade, usando instrumentos de Análise Matemática que alargaram consideravelmente a capacidade de resolver problemas complexos.

Há razões para crer que desde 1783 planeava publicar um tratado sobre a Probabilidade. Como membro da Academia das Ciências de Paris, tinha já escrito um artigo com mais de uma centena de páginas publicado em 1871 na revista daquela instituição, e foi durante um breve período professor de um curso de Probabilidade na Escola Normal Superior, nos efémeros 4 meses que esta durou na sua primeira fundação, num período politicamente conturbado. O seu conhecimento de Probabilidade, e a capacidade de a aplicar (como desde muito cedo tinha feito na área de Astronomia), levaram-no também a ter um papel de relevo na comissão criada para investigar o maior hospital de Paris, pois usou argumentos probablisticos para comparar as taxas de mortalidade nesse hospital com as de outro hospitais franceses e europeus.

A publicação do *Tratado Analítico das Probabilidades* só viria a ocorrer em 1812, com retumbante sucesso, pois houve uma reedição em 1814, consideravelmente aumentada com um longo prefácio que também teve tiragem à parte, com o nome de *Ensaio Filosófico sobre as Probabilidades.* Entre o projecto inicial de 1783 e a sua efectivação em 1812, Laplace tinha inventado instrumentos notáveis como a função geradora de probabilidades, a transformada de Laplace, as funções característica, tinha feito a primeira demonstração geral e rigorosa do Teorema Limite Central, abordara questões de inferência estatística intuindo a importância da probabilidade condicional e estabelecendo regras para o cálculo da probabilidade inversa (porventura desconhecendo a publicação póstuma da obra de Thomas Bayes), e estabelecendo bases rigorosas para o método dos mínimos quadrados e ideias sobre verosimilhança (relacionando claramente este conceito com o de probabilidade *a posteriori*).

Reconhecendo a extraordinária importância do Teorema Limite Central, recomendou a tabulação da distribuição “normal”, e escreveu uma memória fazendo um estudo quase exaustivo da família de distribuições normais. Quase exaustivo, porque lhe escapou uma caracterização da família das normais, que o jovem Gauss estabeleceu em duas páginas magistrais em 1809: é a única família de distribuições absolutamente contínuas, variando em função de um parâmetro de localização, em que a média é o estimador de verosimilhança máxima desse parâmetro de localização. Isto levou Laplace a escrever em 1810 a publicar mais uma “memória”, em que este novo conhecimento já foi incluído e explorado.

No *Ensaio Filosófico sobre a Probabilidade* a discussão tem muitos pontos em comum com a axiomatização de Kolmogorov. Depois de retomar o princípio da razão insuficiente de J. Breroulli para defender o uso de equiprobabilidade dos acontecimentos elementares — o que levou a que se considere que o conceito de probabilidade segundo Laplace é o quociente entre número de casos favoráveis e número de casos possíveis — a discussão é de facto muito mais profunda, e embora a posição dele em inferência estatística seja oscilante entre o que se chama escola frequencista e escola bayesiana, qualquer leitor atento percebe que ele tem uma visão profunda da probabilidade como uma caracterização das nossas convicções em função da informação disponível, e que esta se altera passo a passo, nomeadamente com o uso de probabilidades em cadeia e probabilidades inversas. É pois uma injustiça considerar Laplace um probabilista laplaciano no sentido redutor que foi dado a esta expressão.

A vida pública de Laplace adaptou-se ao regime monárquico, república, império de Napoleão (de quem chegou a ser ministro), refundação da monarquia, sempre rodeado de respeito e honrarias, e nomeado para importantes comissões, nomeadamente naquela que defendeu a adopção do sistema decimal como norma.

Karl Pearson

(1857-03-27 a 1936-04-27) graduou-se em Matemática, no King's College de Cambridge, mas na fase inicial da sua vida profissional interessou-se por uma grande diversidade de outras áreas científicas: foi para a Universidade de Heidelberg para estudar Física e Metafísica, indo seguidamente para a Universidade de Berlin, onde estudou Fisiologia, Direito Romano, Literatura Alemã e Economia Política (foi o seu grande apreço pela obra de Karl Marx que o levou a assinar-se Karl Pearson, sendo o seu nome de baptismo Carl Pearson). Tornou-se um germanista distinto, e o King's College de Cambridge chegou a tentar recrutá-lo para a área de Germânicas.

Apesar de em 1880 já se ter afirmado com publicações importantes nas diversas áreas por que se foi interessando, em 1881 retomou o interesse pela Matemática, primeiro como professor no King's College de Londres, e em 1883 como professor de Matemática (e logo no ano seguinte como professor de Matemática Aplicada) na University College de Londres. Em 1891 foi contratado como professor de Geometria do Gresham College (onde proferiu conferências que ficaram famosas, sendo considerado um excepcional comunicador de ideias), e aí travou amizade com o zoólogo Walter Weldon, que o apresentou a Francis Galton (primo de Charles Darwin), que depressa reconheceu o seu mérito e protegeu a sua carreira. Em 1900, Galton, Weldon e Pearson fundaram a revista *Biometrika*, uma das mais destacadas publicações periódicas de Estatística. A morte prematura de Weldon em 1908 interrompeu uma brilhante parceria de investigação em Biometria e Teoria da Evolução. Em 1911 Galton faleceu, e por disposição testamentária financiou uma cátedra em Eugenismo (posteriormente renomeada Cátedra Galton de Genética), deixando expresso que Pearson deveria ser o primeiro a ocupá-la.

A série de conferências que proferiu no Gresham College no ano em que foi contratado foram a base de um livro notável sobre Filosofia da Ciência, *The Grammar of Science*, diz-se que Albert Einstein o recomendou a amigos e colaboradores como "o primeiro livro que deviam ler". Neste livro (reeditado pela Dover numa edição acessível em 2004) aparecem formas incipientes de Relatividade, e aborda — o que na altura seria ficção científica, se não fosse uma obra tão séria de Filosofia da Ciência — temas como anti-matéria, dobras no tempo, quarta dimensão na descrição física da natureza. Um dos aspectos mais originais de sse livro notável é a defesa da ideia que o paradigma de Ciência estava a mudar da investigação da causalidade para a de quantificação da associação entre fenómenos, uma base para tornar as ciências sociais e humanas tão "respeitáveis"como as ciências experimentais, cujo "determinismo"começava a ser aliás contestado.

Este ponto de vista decerto foi influenciado por Galton, que tinha introduzido as ideias de correlação e de regressão. Pearson de facto usou momentos de produtos dos desvios em abcissas e ordenadas relativamente às respectivas médias para definir rigorosamente e estudar o que muitas vezes se chama coeficiente de correlação de Pearson, mas os conceitos foram primeiro abordados por Galton.

O interesse pelos momentos da distribuições estatísticas — uma ideia que importou da Física — levou-o a caracterizar uma família de distribuições através de parâmetros (com o sentido de "o que está para além das observações") usando funções dos quatro primeiros momentos, média, variância, assimetria e curtose (ou achatamento). Sete tipos de distribuições — incluindo betas, gamas, t de Student como caso limite, e F — surgem como possíveis soluções de uma equação diferencial. Um dos propósitos de K. Pearson era decerto chamar a atenção para distribuições distintas da "normal", mais adequadas para espelhar assimetrias observadas nos dados.

Foi também para mostrar as limitações do modelo normal que Pearson estudou os quadrados dos desvios, publicando em 1900 um influente artigo em que introduz o teste do qui-quadrado (como teste de ajustamento) e a noção de valor de prova ($p - value$). Pearson introduziu muitos termos de uso corrente em estatística, tais como desvio padrão, homocedasticidade e heterocedasticida, e foi pioneiro na apresentação de muitas ideias que estiveram na génese quer da estimação quer dos testes de hipóteses. Este grande homem, porém, não reconheceu a grandeza das ideias de Fisher com quem se incompatibilizou, levando a feíssimos ataques públicos entre eles, que da parte de Fisher perduraram até 30 anos depois de K. Pearson ter falecido! ("A man of more will than intelligence", é como Fisher se refere a K. Pearson no prefácio da última edição dos seus livros em vida dele.)

Há, por outro lado, traços de bom carácter em Pearson, nomeadamente não evitar ser grato, escrevendo e publicando à sua custa uma biografia de Galton, em três volumes, em que claramente assume que Galton foi, de facto, o grande cientista e o catalisador da criação da Estatística moderna.

Depois de ter sido uma figura dominante da Estatística (inclusivamente fundando o primeiro departamento de Estatística do mundo), as áreas que desenvolvera deixaram de ser consideradas prioritárias, e o reconhecimento do seu génio sofreu um eclipse de várias décadas. Actualmente ressurgiu o interesse pelas suas ideias, sendo o seu uso de assimetria e curtose (e as expansões em série de Edgeworth, que foram decerto uma inspiração para as suas ideias) um instrumento importante no estudo de aproximações assintóticas e de métodos de computação intensiva em Estatística Computacional.

Sir Ronal Fisher

Ronald Aylmer Fisher (1890-02-17 a 1962-07-29) é unanimemente considerado o estatístico mais importante de sempre, pelas suas muitas ideias, muitas das quais não só revolucionaram a Estatística, como tiveram um impacte enorme em todas as ciências.

Os seus interesses principais, durante os anos de formação, foram a Matemática e a Biologia Evolutiva. Devido a ter uma incapacidade visual muito forte, a aprendizagem da Matemática foi feita sem usar papel e lápis, o que provavelmente contribuiu para uma capacidade extraordinária para visualizar intelectualmente argumentos geométricos, e por outro lado para estabelecer resultados sem detalhar rigorosamente os passos do seu raciocínio, não se sujeitando à formulação tradicional de apresentar demonstrações incluindo os passos intermédios.

Estudou no Gonville and Caius College, Cambridge, e na rica atmosfera intelectual de Cambridge entusiasmou-se pelas ideias de hereditariedade de Mendel, que ambicionou combinar com as ideias de Darwin sobre evolução. Isso levou-o a fundar, com Maynard Keynes e outros, a Sociedade de Eugenia da Universidade de Cambridge, muito activa, e com intercâmbio muito forte com a Sociedade de Eugenia de Londres, fundada por Francis Galton (que era primo de Darwin).

Os primeiros anos da sua vida profissional, primeiro, durante 6 anos, como estatístico na autarquia de Londres, depois como docente de Física e de Matemática em escolas secundárias, foram materialmente pobres, intelectualmente notáveis. Para além de escrever regularmente críticas de livros científicos para a *Eugenics Review.* Foi ainda nesta fase incipiente da sua investigação que apresentou as primeiras ideias sobre análise da variância (é aliás nesse artigo *The Correlation Between Relatives on the Supposition of Mendelian Inheritance* de 1918, por muitos considerado a fundação da Genética Biométrica, que aparece pela primeira vez a palavra "variância").

Depois de uma fase inicial de boas relações com Karl Pearson, Fisher mostrou que Pearson errara o cálculo do número de graus de liberdade ao aplicar o teste do qui-quadrado na análise de tabelas de contingência. Por outro lado, Pearson recusou artigos que Fisher submetera à *Biometrika*, de que era editor, e deu pareceres negativos a trabalhos que Fihser submetera à Royal Society. Isto originou uma aversão duradoura, de parte a parte, que levou mesmo Fisher a escrever no prefácio da reedição dos seus livros, 30 anos após a morte de Pearson, que aquele era mais "*a man of will than of intelligence*". Apesar das más relações, Karl Pearson ofereceu a Fisher um posto no seu Laboratório de Biometria de Londres, mas Fisher não aceitou o convite, provavelmente por saber que dois homens de de personalidades tão vincadas e antagónicas não poderiam ter uma relação de trabalho estável.

Fisher preferiu por isso aceitar um contrato na Estação Experimental de Rothamsted, onde dispunha de colecções de dados recolhidos ao longo de muitos anos, que lhe permitiram estudos muito completos sobre a variação de colheitas. Entre 1919 e 1925 publicou novos resultados sobre a análise da variância, e foi o pioneiro do desenvolvimento do Planeamento Experimental.

Particularmente importante foi a sua contribuição para a teoria da estimação, com particular ênfase no uso de verosimilhança máxima, e na criação dos conceitos de suficiência e de estatísticas ancilares. Apresentou também uma derivação rigorosa da densidade t de Student, estabelecendo claramente a relação das variáveis qui-quadrado e t com a normal, bem como uma variável conhecida como z de Fisher, para a análise da variância, que viria a ser preterida em favor do quociente de quis-quadrados independentes divididos pelos respectivos números de graus de Liberdade, a F de Fisher-Snedecor.

Em 1925 publicou um dos livros mais influentes em Ciência, *Statistical Methods for Research Workers*, cujas muitas edições foram sendo actualizadas com novas descobertas de Fisher. É difícil fazer uma escolha acertada de quais as principais criações de Fisher, mas é decerto de citar

- Na utilização do seu teste z com dados não normais, recorreu a "testes de aleatorização", inspirando-se em ideias de Peirce (1839-1914), que frutificaram no importante ramo de testes de permutações. Inventou também o teste exacto (hipergeométrico) para a análise de tabelas de contingência duplamente dicotómicas com margens fixas, que é também aplicável para a análise de tabelas cujas frequências esperadas são muito baixas, pondo em causa a aproximação da distribuição exacta (discreta) pela distribuição qui-quadrado.

- Retomou os cumulantes, introduzidos por T. N. Thiele em 1889, mas caídos em desuso, que não têm a interpretação intuitiva dos momentos, mas que se prestam a manipulação algébrica muito mais simples, e que lhe permitiram estender as expansões de Edgeworth, e obter a expansão de Cornish-Fisher que se pode usar para obter aproximaçoes de quantis — nomeadamente do importante indicador *value-at-risk* (*VaR*) em Teoria do Risco e Finanças.

- Introduziu a Informação de Fisher, cujo inverso é o limite inferior de Cramér-Rao da variância de estimadores centrados do parâmetro θ que se pretende estimar.

- Inventou a análise discriminante linear de Fisher, que além das utilizações em áreas mais clássicas da Estatística é usada em Reconhecimento de Padrões e em *Machine Learning*, que identifica uma combinação linear de características que caracteriza classes de objectos ou acontecimento — e consequentemente serve para separar essas classes, servindo para uma redução da dimensionalidade precedendo a classificação.

- Foi um dos introdutores, em 1922, da família exponencial (muitas vezes chamada família exponencial de Darmois-Koopman-Pitman, devido aos trabalhos independentes por eles apresentados por volta de 1935).

- Corrigiu um anterior trabalho de Behrens sobre comparação de valores médios de populações normais quando se dispõe de duas amostras independentes, quando nada se sabe sobre o quociente das respectivas variâncias. Isto viria a despoletar grandes controvérsias em Estatística, com o nome de *problema de Behrens-Fisher*, pois pela primeira vez se constatou que a abordagem bayesiana e a abordagem frequencista podiam levar a soluções diferentes do mesmo problema.

  Bartlett apontou que a solução de Fisher levava estimadores intervalares cuja probabilidade frequencista de cobertura não era a pretendida. De facto, a solução de Fisher coincide com a solução bayesiana, o que não lhe agradava, e o levou a inventar umas "probabilidades fiduciais"que ninguém parece ter entendido, e que foram o insucesso de uma carreira em todos os outros aspectos fulgurante.

A par de toda esta criatividade estatística, Fisher publicou um grande número de trabalhos importantes em Genética Quantitativa.

Após a sua frutuosa carreira em Rothamsted, em 1933 foi contratado como professor de Eugenia na Universidade dee Londres, e em 1943 como professor da Universidade de Cambridge. Forçado a retirar-se por limite de idade, foi para a Universidade de Adelaide, na Austrália, onde manteve intensa actividade de investigação até ao fim da vida.

A vida de Fisher foi ensombrada pelas suas muitas querelas com outros cientistas, como Karl Pearson, Neyman ou Bartlett. Também foi muito criticado pelos seus trabalhos, prventura mercenários, contestando que o tabagismo causa cancro de pulmão. Com algum humor, para estabelecer que correlação não é causalidade, comparou os resultados dos defensores da teoria oposta com um irónico exemplo da correlação entre a importação de maçãs e o aumento de divórcios no Reino Unido .

Para além do já citado *Statistical Methods for Research Workers*, em 1935 publicou novo livro, reportando a sua fundação de *The Design of Experiments*, e em 1956 *Statistical methods and scientific inference*, livro este que é, também, uma reflexão sobre a metodologia da investigação científica.