

ISSN 1677-5473



*Empresa Brasileira de Pesquisa Agropecuária
Secretaria de Gestão e Estratégia
Ministério da Agricultura, Pecuária e Abastecimento*

Texto para Discussão 21

Breve História da Estatística

José Maria Pompeu Memória

Embrapa Informação Tecnológica
Brasília, DF
2004

Exemplares desta publicação podem ser adquiridos na:

Empresa Brasileira de Pesquisa Agropecuária

Edifício-Sede da Embrapa
Secretaria de Gestão e Estratégia
Parque Estação Biológica – PqEB – Av. W3 Norte (final)
CEP 70770-901 – Brasília, DF
Fone (61) 448-4449
Fax: (61) 448-4319

Editor desta série
Antônio Jorge de Oliveira

Coordenador editorial
Vicente G. F. Guedes

Corpo editorial
Antônio Flávio Dias Ávila
Antônio Jorge de Oliveira
Antônio Raphael Teixeira Filho
Ivan Sérgio Freire de Sousa
Levon Yeganiantz

Produção editorial e gráfica
Embrapa Informação Tecnológica

Revisão de texto e tratamento editorial
Francisco C. Martins

Normalização bibliográfica
Dauí Antunes Corrêa

Editoração eletrônica
Júlio César da Silva Delfino

Projeto gráfico
Tênisson Waldow de Souza

1ª edição
1ª impressão (2004): 500 exemplares

Todos os direitos reservados.

A reprodução não autorizada desta publicação, no todo ou em parte,
constitui violação dos direitos autorais (Lei nº 9.610).

Dados Internacionais de Catalogação na Publicação – CIP
Embrapa Informação Tecnológica

Memória, José Maria Pompeu.
Breve história da estatística / José Maria Pompeu Memória. – Brasília, DF :
Embrapa Informação Tecnológica, 2004.
111 p. ; 21 cm. – (Texto para discussão, ISSN 1677-5473 ; 21).

1. Estatística. I. Título. III. Série.

CDD 519.5 (21ª ed.)

© Embrapa 2004

Apresentação

Texto para Discussão é um veículo utilizado pela Secretaria de Gestão e Estratégia, da Empresa Brasileira de Pesquisa Agropecuária – Embrapa –, para dinamizar a circulação de idéias novas e a prática de reflexão e de debate sobre aspectos relacionados à ciência, à tecnologia, ao desenvolvimento agrícola e ao agronegócio.

O objetivo desta série é fazer com que uma comunidade mais ampla, composta de profissionais das diferentes áreas científicas, debata os textos apresentados, contribuindo para o seu aperfeiçoamento.

O leitor poderá apresentar comentários e sugestões, assim como debater diretamente com os autores, em seminários especialmente programados, ou utilizando qualquer um dos endereços fornecidos: eletrônico, fax ou postal.

Os trabalhos para esta coleção devem ser enviados à Embrapa, Secretaria de Gestão e Estratégia, Edifício-Sede, Parque Estação Biológica – PqEB –, Av. W3 Norte (Final), CEP 70770-901 – Brasília, DF. Contatos com a Editoria devem ser feitos pelo fone (61) 448-4449 ou pelo fax (61) 448-4319.

Os usuários da Internet podem acessar as publicações pelo endereço <http://www.embrapa.br/unidades/uc/sge/textdiscussao.htm/>. Para os usuários do Sistema Embrapa, basta clicar em novidades, na Intranet.

O Editor

Sumário

Introdução	9
Primórdios	11
A Escola Biométrica	26
A Fase da Experimentação	36
Desenvolvimento dos Levantamentos por Amostragem	68
A Era Atual	81
Referências	92

21



Breve História da Estatística

José Maria Pompeu Memória¹

¹ Pesquisador aposentado da Empresa Brasileira de Pesquisa Agropecuária – Embrapa – Parque Estação Biológica – PqEB
Av. W3 Norte (final), CEP 70770-900 Brasília, DF, Brasil

Introdução

*Auguste Comte*¹

“Não se conhece completamente uma ciência,
a menos que se saiba sua história”.



esta *Breve História da Estatística*, o assunto está exposto didaticamente em cinco seções, a saber:

- *Primórdios*, dividida em três partes: Contribuição da Astronomia, A Influência de Quételet, e As Idéias de Galton.
- *A Escola Biométrica*.
- *A Fase da Experimentação*, também dividida em três partes: Fundamentos Teóricos da Estatística, Análise de Variância e Delineamentos Experimentais, e Teste de Hipóteses Estatísticas.
- *Desenvolvimento dos Levantamentos por Amostragem* e, finalmente.
- *A Era Atual*, onde é enfatizada a revolução causada pelos computadores.

¹ Do *Cours de Philosophie Positive*. Citação extraída do livro *Histoire de la Démographie* de Jacques et Michel Dupâquier – Paris: Librairie Academique Perrin, 1985. p. 17.

Contudo, a exposição não tem caráter puramente descritivo das descobertas cronologicamente acumuladas, mas procura salientar as influências recíprocas entre as personalidades mais representativas e suas idéias sobre a origem e o desenvolvimento dos métodos e técnicas mais importantes, que contribuíram na formação da Estatística como disciplina autônoma, numa perspectiva integrada à própria história da ciência.

Não é tarefa fácil saber quando se originou a história de qualquer ramo do conhecimento, pois isso vai depender do conceito que fizermos dele e que, naturalmente, evoluirá no decorrer do tempo. A história da Estatística bem confirma esta asserção.

Ainda hoje, no conceito popular a palavra estatística evoca dados numéricos apresentados em quadros ou gráficos, publicados por agências governamentais, referentes a fatos demográficos ou econômicos. A etimologia da palavra, do latim *status* (estado), usada aqui para designar a coleta e a apresentação de dados quantitativos de interesse do Estado, bem reflete essa origem. Entretanto, a mera coleta de dados assim apresentados está longe de ser o que entendemos, hoje, por Estatística. Na verdade, sua feição essencial é a de ser um conjunto de métodos (métodos estatísticos), especialmente apropriado, no dizer de George Udny Yule (1871 – 1951), ao tratamento de dados numéricos afetados por uma multiplicidade de causas. Esses métodos fazem uso da Matemática, particularmente do cálculo de probabilidades, na coleta, apresentação, análise e interpretação de dados quantitativos.

Primórdios



Desde remota antiguidade, os governos têm se interessado por informações sobre suas populações e riquezas, tendo em vista, principalmente, fins militares e tributários. O registro de informações perde-se no tempo. Confúcio relatou levantamentos feitos na China, há mais de 2000 anos antes da era cristã. No antigo Egito, os faraós fizeram uso sistemático de informações de caráter estatístico, conforme evidenciaram pesquisas arqueológicas. Desses registros também se utilizaram as civilizações pré-colombianas dos maias, astecas e incas. É conhecido de todos os cristãos o recenseamento dos judeus, ordenado pelo Imperador Augusto.

Os balancetes do império romano, o inventário das posses de Carlos Magno, o *Doomsday Book*, registro que Guilherme, o Conquistador, invasor normando da Inglaterra, no século 11, mandou levantar das propriedades rurais dos conquistados anglo-saxões para se inteirar de suas riquezas, são alguns exemplos anteriores à emergência da estatística descritiva no século 16, na Itália.

Essa prática tem sido continuada nos tempos modernos, por meio dos recenseamentos, dos quais temos um exemplo naquele que se efetua a cada decênio, em nosso País, pela Fundação IBGE, órgão responsável por nossas estatísticas (dados estatísticos) oficiais.

Com o Renascimento, foi despertado o interesse pela coleta de dados estatísticos, principalmente por suas aplicações na administração pública. A obra pioneira de Francesco Sansovini (1521 – 1586), representante da

orientação descritiva dos estatísticos italianos, publicada em 1561, é um exemplo dessa época. Deve ser mencionado ainda o reconhecimento por parte da Igreja Católica Romana da importância dos registros de batismos, casamentos e óbitos, tornados compulsórios a partir do Concílio de Trento (1545 – 1563).

Entretanto, mais amplos e gerais foram os estudos feitos pelos alemães, especialmente por Gottfried Achenwall (1719 – 1772), professor da Universidade de Göttingen, a quem se atribui ter criado o vocábulo estatística, em 1746. Contudo, nada mais fizeram do que dar melhor sistematização e definição da mesma orientação descritiva dos estatísticos italianos.

Acreditar nessas atividades como o começo da história da estatística é deixar de compreender o verdadeiro significado da Estatística. Podemos dizer que o desenvolvimento da estatística teve origem nas aplicações, pois nenhuma disciplina tem interagido tanto com as demais disciplinas em suas atividades do que ela, dado que é por sua natureza a ciência do significado e do uso dos dados. Daí, sua importância como instrumento auxiliar na pesquisa científica.

A primeira tentativa para se tirar conclusões a partir de dados numéricos foi feita somente no século 17, na Inglaterra, com o que foi denominado *Aritmética Política*, que evoluiu para o que se chama hoje de demografia. Contudo, só começou realmente a existir como disciplina autônoma no raiar do século 20, o verdadeiro início da estatística moderna.

A tentativa acima referida foi feita por John Graunt (1620 – 1674), um próspero negociante londrino de tecidos que em 1662, publicou um pequeno livro intitulado

Natural and Political Observations Mentioned in a Following Index and Made upon the Bills of Mortality. Sua análise foi baseada sobre razões e proporções de fatos vitais, nos quais ele observou uma regularidade estatística num grande número de dados. Por seu trabalho foi eleito *Fellow of the Royal Society (F. R. S.)*, sociedade científica fundada em 1660, por Carlos II.

Os dados usados por Graunt compreendiam uma série anual de 1604 a 1660, coletados nas paróquias de Londres, de onde ele tirou as seguintes conclusões: que havia maior nascimento de crianças do sexo masculino, mas havia distribuição aproximadamente igual de ambos os sexos na população geral; alta mortalidade nos primeiros anos de vida; maior mortalidade nas zonas urbanas em relação às zonas rurais.

Graunt era cômico de ser leigo no assunto, pois não era médico, nem matemático, nem político, mas apenas uma mente curiosa que utilizou com lógica uma análise, pode-se dizer “científica”, dos registros disponíveis sobre mortalidade. Com seus dados, elaborou uma tábua de vida rudimentar, baseada apenas na sobrevivência nas idades de 6 a 76 anos.

Foi William Petty (1623 – 1683), contemporâneo e continuador de Graunt, quem denominou de Aritmética Política à nova arte de raciocinar por meio de dados sobre fatos relacionados com o governo. Em 1683, ele publicou sua obra *Five Essays on Political Arithmetic* e sugeriu que fosse criada uma repartição de registro de estatísticas vitais, mas isso só se consolidou no século 19, com o Dr. William Farr (1807 – 1883), contribuidor original da estatística médica. Note-se que a denominação

posterior de estatística acabou por incluir a Estatística Descritiva e a Aritmética Política.

Dos trabalhos desse período, sem dúvida, o mais importante foi o do astrônomo inglês Edmond Halley (1656 – 1742), que em 1693 construiu a primeira tábua de sobrevivência, elaborada com os registros vitais da cidade alemã de Bresláu (atual Wroclaw, Polônia), referentes ao período de 1687 a 1691, elemento básico para o cálculo de seguros de vida. Embora o seguro comercial tivesse sido praticado pelos babilônios e fosse conhecido dos gregos e dos romanos, Halley é, com justiça, considerado o criador do cálculo atuarial. Deve ser ainda mencionado o nome de Richard Price (1723 – 1791), o fundador da atuária, na Inglaterra.

Contemporâneo desse período em que as idéias estatísticas tiveram início, desenvolveu-se o cálculo de probabilidades, mas independentemente dessas idéias, vindo entretanto a influenciá-las posteriormente. O cálculo de probabilidades originou-se da correspondência entre dois grandes matemáticos do século 17: Blaise Pascal (1623 – 1662) e Pierre de Fermat (1601 – 1665), para solucionar problemas relacionados com jogos de azar, em moda nos salões da França, sustentados pelo lazer de uma aristocracia. Desses problemas, os mais célebres foram propostos a Pascal em 1654, pelo nobre francês Chevalier de Méré, jogador de grande experiência e perspicácia.

Na verdade, antes de Pascal e Fermat, já alguns matemáticos italianos como Niccolò Fontana Tartaglia (1499 – 1557), Girolamo Cardano (1501 – 1576), seguidos por Galileu Galilei (1564 – 1642) interessaram-se por problemas de probabilidades relacionados com jogos de dados.

Os primeiros problemas sobre probabilidades refletiram o desenvolvimento da análise combinatória em jogos de azar. Em todos eles eram examinados os diferentes modos em que arranjos e combinações podiam ser empregados na enumeração dos casos favoráveis. Esses problemas eram dominados por considerações sobre os casos igualmente prováveis, com as probabilidades determinadas a priori, onde foi utilizado o seguinte tipo de raciocínio: dado uma urna contendo a (bolas pretas) e b (bolas brancas), a probabilidade de se extrair uma bola preta é igual a $\frac{a}{a+b}$.

O primeiro matemático a considerar situações em que não era possível a enumeração de casos igualmente possíveis foi Jacob Bernoulli (1654 – 1705), professor da Universidade de Basileia, Suíça, e primeiro membro de uma numerosa família de matemáticos suíços, que propôs determinar a probabilidade de tais casos a posteriori, isto é, pela frequência relativa do número de casos favoráveis determinada empiricamente, em sua obra *Ars Conjectandi*, publicada postumamente em 1713, por seu sobrinho Nicholas Bernoulli.

A novidade consistia na tentativa de dar um tratamento formal à noção vaga de quanto maior fosse o acúmulo de evidência sobre uma desconhecida proporção de casos, mais próximo estar-se-ia de aceitar essa proporção, isto é, à medida que o número de observações aumenta. Pode-se afirmar que J. Bernoulli abriu o caminho para a quantificação da incerteza com seu teorema cujo enunciado, na sua forma moderna, é o que se conhece como a *lei fraca dos grandes números*. Numa de suas formas mais conhecidas, pode ser expresso da seguinte maneira:

$$P_{N \rightarrow \infty} \left(\left| \frac{X}{N} - p \right| < \varepsilon \right) \rightarrow 1 ,$$

onde a probabilidade P tende para a certeza, quando o número de observações N cresce indefinidamente, sendo X o número de sucessos observados, p a proporção desconhecida e $\varepsilon > 0$, um número tão pequeno quanto se queira. Este teorema serve de base para justificar a noção intuitiva de probabilidade medida pela frequência relativa:

$$\lim_{N \rightarrow \infty} \frac{X}{N} = p .$$

Bernoulli tratou de número de sucessos, tendo lidado naturalmente com a distribuição binomial.

O primeiro passo que se seguiu coube a um exilado francês huguenote (nome dado aos protestantes expatriados da França por ocasião da revogação do edital de Nantes, em (1685), que vivia em Londres, Abraham De Moivre (1667–1754), famoso por sua obra, em várias edições, *The Doctrine of Chance* (1718, 1738, 1756). Sua maior contribuição foi a de chegar à curva normal como limite da distribuição binomial, em artigo publicado em 1733, tendo se utilizado da expressão $n!$ para n grande, atribuída a Stirling, mas que ele derivou primeiramente:

$$n! = \left(\frac{n}{e} \right)^n \sqrt{2 \pi n} \text{ (aprox.)}$$

Paralelamente ao trabalho dos probabilistas, desenvolveram-se métodos de grande utilidade no tratamento dos dados de observação, em particular da Astronomia e da Geodésia, de onde surgiu a *Teoria dos Erros*. A importância da curva normal e o uso amplo da palavra erro, bem demonstram o quanto desses conceitos foi incorporado à teoria estatística, o que justifica a abertura de uma parte sobre essa contribuição.

Contribuição da Astronomia

De há muito tempo que os astrônomos tinham soluções práticas para lidar com o problema de conciliar observações discordantes como, por exemplo, tomando a média aritmética dessas observações, após descarte daquelas muito discordantes (*outliers*). Entretanto, foi somente no século 18 que ocorreu nos de mente mais matemática dar um tratamento que conduziu à curva de frequência dos erros acidentais. Assim, os trabalhos mais importantes devem-se a dois dos maiores matemáticos de todos os tempos: Pierre Simon, Marquês de Laplace (1749 – 1827) e Carl Friedrich Gauss (1777 – 1855).

A maior contribuição de Laplace, na teoria de probabilidades, é hoje conhecida por teorema central (fundamental) do limite e pode ser descrita como uma generalização do teorema do limite de De Moivre. Na sua forma clássica, o *Teorema Central do Limite* enuncia que:

“Qualquer soma ou média de variáveis aleatórias tem, para um grande número de termos, uma distribuição aproximadamente normal”.

Em 1810, Laplace publicou *Mémoire sur les formules qui sont fonction de très-grands nombres*. Vários eminentes matemáticos contribuíram, posteriormente, para seu desenvolvimento, emprestando o maior rigor. Atualmente, podemos enunciá-lo em sua forma clássica, do seguinte modo: se uma população tem variância finita σ^2 e média μ , a distribuição da média aritmética de n observações aproxima-se de uma distribuição normal com variância $\frac{\sigma^2}{n}$ e média μ , à medida que o tamanho n da amostra aumenta. O fato notável é que nenhuma restrição é feita sobre a função de densidade da população original. Daí, a grande importância da distribuição normal na teoria e aplicações da Estatística. As contribuições de Laplace foram sintetizadas na sua obra monumental *Théorie Analytique des Probabilités*, publicada em 1812.

Independentemente, Gauss chegou à curva dos erros com espírito empírico, adotando como axioma o princípio de que o valor mais provável de uma quantidade desconhecida, observada com igual precisão várias vezes sob as mesmas circunstâncias, é a média aritmética das observações.

Após várias considerações gerais a priori sobre a função de frequência dos erros $f(x)$, como a de ter um máximo para $x = 0$, de ser simétrica e admitir valor nulo fora do limite dos erros possíveis, obteve a curva dos erros, que ficou conhecida como a curva de Gauss. Seu trabalho foi publicado em 1809, com o título latino *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solum Ambientium*, com uma seção final sobre combinações

das observações e a curva dos erros. Esses estudos levaram-no a enunciar o *Princípio dos Mínimos Quadrados*.

Com efeito, numa distribuição normal, a probabilidade dos erros x_1, x_2, \dots, x_n ocorrerem simultaneamente é máxima, quando a soma dos quadrados dos desvios ou erros for mínima. O valor obtido para que essa condição se verifique é a média aritmética das observações, que fica assim justificada como o melhor meio de se combinar as observações, se os erros seguem uma distribuição normal. O argumento de Gauss pode ser acusado de circular.

Posteriormente, Gauss melhorou sua argumentação, publicando em 1823, a obra *Theoria Combinationis Observationum Erroribus Minimis Obnoxiae*. Entretanto, justiça deve ser feita ao matemático francês André Marie Legendre (1752 – 1833), que em 1805 já havia proposto o método dos mínimos quadrados como aplicável para combinar observações astronômicas e geodésicas, baseado em critério intuitivo. Gauss, no entanto, considerou ter prioridade na idéia, pois já vinha se preocupando com ela desde 1795. A obra de Legendre, chamada *Nouvelles Méthodes pour la Détermination des Orbites des Comètes*” continha o apêndice *Sur la méthode des moindres quarrés*.

A curva normal ficou mais conhecida como curva de Gauss e o princípio dos mínimos quadrados foi atribuído a Gauss, que o utilizou como método de estimação de parâmetros. Pode-se invocar aqui, a Lei de Eponímia de Stigler, citada em Stigler (1980, p. 147), “Nenhuma descoberta científica é conhecida pelo nome do seu inventor original”.

Em seguida, serão vistas outras influências na formação da Estatística, a corrente social e a corrente biológica, representadas respectivamente pelos grandes vultos de Quételet e Galton.

A influência de Quételet

Adolphe Quételet foi quem primeiro percebeu que a Estatística deveria ser baseada na noção de probabilidade. Ninguém, melhor do que ele, representa a nova influência oriunda das ciências sociais (chamadas, na época, de “morais”), trazendo de volta a preocupação com o social originada pela Escola de Aritmética Política.

De cultura universal, Quételet foi matemático, astrônomo, físico, estatístico, poeta, escritor e até artista. Lambert Adolphe Jacques Quételet (1796–1874) nasceu em Gand e faleceu em Bruxelas, Bélgica. Com justiça, é considerado o “pai das estatísticas públicas”, e o iniciador da colaboração internacional. Suas duas maiores contribuições na análise estatística dos dados sociais foram o conceito de homem médio e o ajustamento da distribuição normal, conjugados com a interpretação da regularidade estatística. Sua principal obra, *Essai de Physique Sociale*, publicada em 1835, é dividida em quatro volumes. Os dois primeiros são consagrados às qualidades físicas do homem, o terceiro às qualidades morais e intelectuais, e o quarto sobre as propriedades do homem médio e o sistema social. Seu conceito de homem médio é uma ficção estatística destinada a facilitar as comparações no espaço e no tempo. Neste conceito, a originalidade de Quételet não é ter calculado as médias aritméticas das medidas, mas em ter considerado suas

dispersões e descoberto que a curva normal (este nome foi dado, posteriormente, por Galton e Pearson à curva dos erros, como era então conhecida) podia ser ajustada satisfatoriamente às medidas de peso, estatura e perímetro torácico por ele feitas em recrutas franceses. Coletou, também, dados sobre criminalidade e delinquência, agrupando-os de acordo com o sexo, a idade, a escolaridade e o tipo de delito, introduzindo a idéia de predisposição ao crime.

Quételet foi um organizador eficiente e hábil. São exemplos desse traço de sua personalidade, a construção do Observatório Astronômico de Bruxelas, a criação da *Comission Centrale de Statistique*, em 1841, e a organização do *Congresso Internacional de Meteorologia*, em 1873, em Viena, pouco antes de falecer. Sua grande idéia era a constituição de uma rede internacional de observações meteorológicas, segundo métodos uniformes. Planejou um recenseamento do Reino dos Países Baixos para 1829 (que então compreendia Bélgica, Holanda e Luxemburgo), tendo pensado em utilizar o método da estimativa da razão para estimar a população, conhecido por método do multiplicador, devido a Laplace, que o empregou em 1802, determinando o número de nascimentos na França e multiplicando esse número pela razão da população para os nascimentos. Essa razão fora estimada tomando-se o censo em algumas comunidades selecionadas, escolhidas em 30 departamentos (distritos) espalhados por todo o país, e o número de nascimentos nessas comunidades baseado na média dos 3 anos 1800, 1801 e 1802.

Laplace não tinha idéia de amostra aleatória, mas usou um processo que se aproximou dela, tendo obtido para a mencionada razão, o valor 28,35282, usado como

multiplicador. Entretanto, Quételet não se utilizou desse critério, por achar que havia muita variação nos Países Baixos, o que demandaria um grande número de regiões a serem escolhidas para sua determinação, sendo tão dispendioso quanto se fosse feita a enumeração total. Essa enumeração acabou sendo feita em 1846, no recenseamento geral da população, agricultura e indústria para a Bélgica, já então independente desde 1830.

As atividades internacionais de Quételet não foram menos notáveis. Assim, em 1853, organizou o primeiro Congresso Internacional de Estatística, em Bruxelas, iniciativa que em 1885, levou à criação do Instituto Internacional de Estatística, em Londres. A sede atual desse Instituto é em Haia, na Holanda. *Quételet* foi também responsável pela fundação da Statistical Society of London, em 1834, posteriormente denominada Royal Statistical Society.

As idéias de Galton

Sir Francis Galton (1822 – 1911) é considerado por Stigler (1986), um personagem romântico na História da Estatística e talvez o último dos cientistas fidalgos.

Foi a leitura do livro de Charles Darwin (de quem era meio primo em primeiro grau) sobre *Origin of Species* (1859), responsável em transformá-lo de geógrafo amador em antropólogo e eugenista (a palavra eugenia foi cunhada por ele, em 1883).

Sob o ponto de vista estatístico, seu livro *Natural Inheritance*, publicado em 1889, é provavelmente sua obra mais importante. Ele elaborou a sugestão de que a distribuição normal é completamente determinada pela

mediana e o desvio semi-quartilico, tendo usado preferencialmente a distribuição cumulativa de frequência, à qual cognominou de ogiva. Inventou, também, um engenhoso aparelho chamado quincunx para ilustrar como poderia ser visualizada a curva normal.

Entretanto, as mais notáveis contribuições de Galton foram a enumeração explícita e parcialmente quantitativa dos conceitos de regressão e correlação. Em suas pesquisas sobre o diâmetro das sementes de ervilha-de-cheiro, chegou a uma equação, por ele denominada de reversão à média:

$$y - \bar{y} = r(x - \bar{x}),$$

onde os y exprimem o diâmetro das sementes das plantas-filhas e os x o das plantas-pais, r significando reversão, posteriormente chamada de regressão. Contudo, foi no estudo comparativo da estatura entre pais e filhos, em 1885, que Francis Galton usou, pela primeira vez, o termo regressão, para denotar a regressão à média da população por ele observada, pois quando os pais eram mais altos do que a média, os filhos tendiam a ser menores do que eles e, quando os pais eram mais baixos que a média, os filhos tendiam a ser maiores do que eles.

Na tábua de frequência bidimensional, representada pela altura dos filhos adultos e dos meio-pais (média do pai e da mãe – a estatura das mulheres foi multiplicada por 1,08), Galton observou que os contornos de igual frequência eram constituídos por elipses concêntricas semelhantemente dispostas e traçou as linhas de regressão à mão, tendo achado que a declividade da linha de regressão dos pais em relação à dos filhos era metade da

declividade da linha de regressão dos filhos em relação aos pais, uma vez que o meio-pai era a média de duas alturas não correlacionadas, cada uma com a mesma dispersão populacional.

Os conhecimentos matemáticos de Galton não eram suficientes para lidar com o problema e então ele apelou para seu amigo J. D. Hamilton Dickson, professor de Matemática na Universidade de Cambridge, para achar a fórmula da superfície encontrada, que nos tempos atuais corresponde à da função normal bidimensional.

Ainda não havia ocorrido a Galton a expressão exata da correlação, pois desconhecia seu sinal. Ele expressou-se sobre a co-relação (só depois escrita correlação), como consequência das variações devidas a causas comuns. A letra r foi inspirada na letra inicialmente usada para designar a reversão. Os efeitos de posição e de escala das observações das variáveis foram eliminados com a padronização das variáveis por meio da centragem sobre a mediana e pela eliminação do efeito escala pela divisão pelo desvio semiquartilico. Contudo, essa padronização trazia a inconveniência de produzir valores de r maiores que a unidade.

A fórmula por ele proposta foi modificada por Walter Frank Raphael Weldon (1860 – 1906), professor de Zoologia em Cambridge, muito ligado a Galton, que chegou à necessidade de se atribuir um sinal positivo ou negativo ao examinar as dimensões de 22 pares de órgãos de várias espécies de camarões. Entretanto, a fórmula do coeficiente de correlação, como é hoje conhecida, só foi determinada em 1896, por Karl Pearson.

Galton interessou-se, também, pela utilização dos métodos estatísticos em Psicologia. Conjuntamente com

o psicólogo americano James McKeen Cattell (1860 – 1944), pioneiro desses estudos nos Estados Unidos, produziu dados estatísticos de distribuição assimétrica que contrariavam suas concepções teóricas, nas quais predominavam o conceito de curva normal. Mais uma vez, venceu sua limitação em matemática com a ajuda de seu amigo matemático Donald MacAlister, que sugeriu a transformação dos dados por seus logaritmos, surgindo, daí, a distribuição normal logarítmica, usando a média geométrica como medida de posição mais apropriada.

Os interesses de Galton foram amplos e variados, bem próprios da versatilidade de seu talento, mas na Psicologia, sua preocupação precípua foi a medida das diferenças individuais e a influência desempenhada pela hereditariedade (*nature*) e pelo meio ambiente (*nurture*) na manifestação dessas diferenças, tanto nos traços físicos como nos mentais, particularmente na inteligência.

Galton foi herdeiro de grande fortuna, tendo deixado filantropicamente 45 mil libras esterlinas para a fundação de uma cátedra de Eugenia, na Universidade de Londres e, 200 libras esterlinas para a publicação de uma revista destinada às aplicações da Estatística à herança biológica, denominada *Biometrika*, cujo primeiro número saiu em 1901, e é publicada até hoje, sem interrupção, embora com a finalidade ampliada.

Em 1904, fundou um laboratório de pesquisas sobre a genética humana, que veio a ser denominado Galton Laboratory of National Eugenics cujo primeiro diretor foi Karl Pearson, que com o próprio Galton e foram os fundadores da Escola Biométrica. Em reconhecimento as suas inúmeras contribuições, foi-lhe concedido o título honorífico de Sir, em 1909.

A Escola Biométrica



Escola Biométrica floresceu na Inglaterra, entre o final do século 19 e o começo do século 20, mais precisamente entre 1890 e 1920. Foi um dos grandes períodos formativos da história da Estatística, com a predominância das técnicas de correlação e ajustamento de curvas, de notáveis resultados na descrição das grandes amostras. Seu principal representante foi Karl Pearson (1857 – 1936), considerado, com justiça, o fundador da Estatística.

Pearson estudou Matemática em Cambridge, a universidade britânica de maior tradição nesse assunto. Antes de se interessar pela estatística, exerceu várias atividades, entre elas a de lecionar matemática aplicada e mecânica no curso de engenharia do University College – nome dado à instituição de ensino superior do antigo sistema de educação britânico afiliado à Universidade de Londres. Homem de grande erudição, capacidade de trabalho e forte personalidade, seu pensamento filosófico influenciou suas idéias estatísticas.

Em 1892, publicou sua obra clássica de filosofia, *The Grammar of Science*, de grande repercussão. De acordo com Pearson, toda variação se dava numa escala contínua; as variáveis descontínuas ou discretas seriam variáveis contínuas com interrupções, e as medidas de associações entre elas teriam o propósito de, na verdade, estimar a correlação subjacente entre as variáveis contínuas. Essas pressuposições foram feitas porque ele estava convencido de que a concepção unificada da ciência era possível graças ao conceito de correlação no lugar da

casualidade, sendo esta considerada como limite teórico da correlação perfeita. Essa metafísica de Pearson teve considerável influência na Psicometria, pois para muitos, somente a variação contínua era mensurável e, portanto, a suposição de uma variação contínua subjacente às características estudadas elevaria a Psicologia ao nível das ciências experimentais, das quais a Física era o modelo por excelência.

Isso mostra como foi difícil para Pearson aceitar a herança mendeliana, de caráter essencialmente descontínuo, como de discórdia com os biometristas, para os quais a variação fenotípica contínua seria incompatível com a variação genotípica descontínua. Coube, a R. A. Fisher (1918), mostrar que as correlações entre parentes estudadas pelos biometristas, representados por Weldon e Pearson, eram explicadas dentro da hipótese mendeliana da hereditariedade.

Influenciado pelas idéias de Galton e de seu amigo Weldon, Pearson inicialmente absorveu-se em problemas de estatística relacionados com a herança biológica numa série de artigos intitulados *Mathematical Contributions to the Theory of Evolution*, de 1889 a 1912, a maioria publicada em *Biometrika*, da qual foi editor de 1901 até sua morte, em 1936. Em seu artigo *Regression, Heredity and Panmixia*, Pearson (1896) propôs, para o coeficiente de correlação, a fórmula do momento-produto, tal qual é conhecida hoje:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}},$$

Nessa fórmula, usou os desvios em relação à média aritmética e não à mediana, e o desvio-padrão (termo por ele cunhado), em vez do desvio semiquartilico, conforme tinha sido usado anteriormente.

Seus estudos subseqüentes levaram-no ao desenvolvimento da regressão e correlação múltiplas e foram a base de numerosos artigos escritos pelo próprio Pearson e seus colaboradores, distinguindo-se entre eles George Udny Yule, que foi seu assistente de 1893 a 1912.

Yule é conhecido, especialmente, por seu livro *An Introduction to the Theory of Statistics*, cuja primeira edição é de 1911 e, a partir da décima-primeira, em 1937, contou com a co-autoria do renomado estatístico Maurice George Kendall (1907 – 1983). Esse livro de Yule & Kendall (1937) foi por muito tempo o texto clássico sobre métodos estatísticos de maior aceitação, principalmente entre os economistas. Em 1912, Yule transferiu-se para Cambridge, como lecturer (professor-assistente), posteriormente promovido a reader (professor-adjunto), cargo em que se aposentou precocemente, em 1931.

Em analogia com o coeficiente de correlação como medida da associação de dados quantitativos, Yule introduziu o coeficiente de associação para a medida do grau de associação em tabelas de contingência 2 x 2, que levou a longa controvérsia quando foi sugerido considerar-se a tabela 2 x 2 como dupla dicotomia de uma superfície normal bidimensional. Karl Pearson (1901) desenvolveu um método de estimar essa correlação, chamando-a de correlação tetracórica. Hoje, tanto o coeficiente de associação como a correlação tetracórica caíram em desuso, sendo considerados apenas uma curiosidade estatística. Em seus trabalhos sobre correlação e regressão

seriais, Yule inventou o correlograma e as séries autorregressivas. Ele também abriu o caminho para a derivação da distribuição amostral do coeficiente de correlação parcial, considerando-a da mesma forma que a da correlação total. Contudo, foi Fisher (1924) quem resolveu essa derivação.

Além de sua contribuição na teoria da correlação, Pearson interessou-se, também, pelas distribuições de frequência que não seguiam a curva normal, com a intenção de agrupá-las todas, subordinando-as a um sistema geral de curvas de frequência, conhecido por Sistema de Curvas de Pearson. Ele tinha em mente a unificação desse sistema por meio da estimação dos parâmetros pelo Método dos Momentos. O ajustamento dessas curvas e a comprovação da aderência desses ajustamentos levou-o a descobrir e utilizar o teste do χ^2 (qui-quadrado) como teste de significância, embora usando o número errado de graus de liberdade, corrigido, posteriormente, por R. A. Fisher.

Tendo observado que a curva normal tinha a mesma declividade nos diversos pontos iguais à da distribuição binomial, Pearson interessou-se em saber que curva teria essa propriedade para o caso da distribuição hipergeométrica, pressupondo que seria uma curva de frequência generalizada. Segundo ele, uma classe geral de distribuições seria obtida pelas soluções da equação diferencial:

$$\frac{1}{y} \cdot \frac{dy}{dx} = \frac{x+a}{b_0 + b_1x + b_2x^2},$$

onde $\frac{dy}{dx}$ é a declividade da reta que liga dois pontos sucessivos da distribuição hipergeométrica. Assim,

diferentes famílias de curvas são obtidas quando se admitem várias relações entre as constantes a , b_0 , b_1 e b_2 . As soluções dessa equação foram classificadas, por Pearson, em 12 famílias de curvas, aquelas de determinada família, sendo chamadas do Tipo I, as de uma segunda família do Tipo II, e assim, sucessivamente; a curva normal é do Tipo VII e é obtida pela integração da citada equação diferencial, fazendo b_1 e $b_2 = 0$. As diversas curvas de Pearson têm várias formas causadas por um grau maior ou menor de assimetria e achatamento. Foram largamente usadas nas amostras grandes, mas, atualmente, apresentam apenas interesse acadêmico. O artigo fundamental de Pearson (1895) sobre esse assunto foi publicado em *Contributions to the Mathematical Theory of Evolution*. A quem se interessar por maiores detalhes, aconselha-se consultar o livro de W. P. Elderton (1927).

O conhecimento dos índices de assimetria e de achatamento ou curtose (kurtosis), conforme denominado por Pearson (cuja preferência pelos termos de origem grega era notória), requer os conhecimentos de certas quantidades obtidas por meio dos chamados momentos de ordem superior. Pearson tomou o termo momento emprestado da Física e fez largo uso desse conceito, a partir de 1893, mas já era conhecido antes.

No Sistema de Pearson, o conjunto de funções contínuas tem seus parâmetros expressos em termos de momentos; as curvas ficarão determinadas, se conhecermos os quatro primeiros momentos. O momento de ordem k em relação à média μ é o valor esperado $\mu_k = E(x - \mu)^k$. Assim, para $k = 2$ obtém-se a variância $\mu_2 = E(x - \mu)^2 = \sigma^2$.

Usando-se a variável reduzida $u = (x - \mu)/\sigma$ obtém-se $E(u^k) = \alpha_k$, sendo

$$\alpha_1 = 0,$$

$$\alpha_2 = 1,$$

$$\alpha_3 = \frac{E(x - \mu)^3}{\sigma^3} = \frac{\mu_3}{\sigma^3},$$

$$\alpha_4 = \frac{E(x - \mu)^4}{\sigma^4} = \frac{\mu_4}{\sigma^4},$$

onde α_3 e α_4 são números abstratos usados, respectivamente, como medidas da assimetria e da curtose. Numa distribuição simétrica $\alpha_3 = 0$, se $\alpha_3 > 0$ a assimetria é positiva, a curva é enviesada à esquerda, com longa cauda à direita; se $\alpha_3 < 0$ a assimetria é negativa, a curva é enviesada à direita, com longa cauda à esquerda. Para $\alpha_4 = 3$, caso da curva normal, a distribuição é mesocúrtica; $\alpha_4 > 3$ é leptocúrtica, alongada, com caudas longas dos dois lados; $\alpha_4 < 3$ é platicúrtica, achatada no alto.

O método dos momentos, como método de estimação dos parâmetros, foi uma das grandes contribuições de Pearson na inferência estatística. As estimativas dos parâmetros são calculadas com dados de grandes amostras, na pressuposição de que elas são boas estimativas e, como o erro de amostragem é mínimo, os momentos populacionais são obtidos igualando-os aos momentos amostrais.

Outra grande contribuição de K. Pearson à inferência estatística foi o uso do χ^2 como teste de

significância para comprovar o ajustamento das curvas de frequência (goodness of fit). Para isso, Pearson (1900) usou a distribuição do qui-quadrado por ele descoberta. Na verdade, foi uma redescoberta, embora Pearson não tivesse conhecimento, pois o astrônomo alemão F. R. Helmert já a havia descoberto em 1875, mas noutro contexto, constituindo mais um exemplo da lei de eponímia, de Stigler. Entretanto, nas aplicações do teste do qui-quadrado tanto no teste de aderência (goodness of fit) das curvas de frequência, como na existência de associação nas tabelas de contingência, usou o número errado de graus de liberdade, por não ter percebido corretamente as restrições impostas e a perda de um grau de liberdade para cada parâmetro estimado. Assim, é que no teste de significância da existência de associação, numa tabela 2×2 , Pearson usou o teste com três graus de liberdade, em vez de um grau de liberdade. Yule percebeu esse erro e conseguiu mostrar, construindo um grande número de tabelas, que para uma tabela de contingência $r \times c$, o número correto de graus de liberdade seria $(r-1)(c-1)$, mas não conseguiu provar sua proposição, restando a Fisher (1922a) fazê-la.

Com o aumento de trabalho, Pearson procurou angariar recursos para continuar as atividades do Laboratório de Biometria, obtendo em 1903, a primeira de uma série de doações. Uma doação de Galton levou à fundação do Laboratório de Eugenia, em 1907. Esses laboratórios foram combinados em 1911, quando foi estabelecido o Departamento de Estatística Aplicada, e Pearson passou a ocupar a cátedra Galton de Eugenia no University College, cargo em que permaneceu até sua aposentadoria, em 1933.

De 1894 a 1930, o University College, em Londres, foi o único lugar no Reino Unido para ensino avançado de Estatística, para onde afluíam estudantes do país e do exterior, para pós-graduação. Foi nessa época que Karl Pearson, o “sumo pontífice” de então em matéria de estatística, foi procurado por William Sealy Gosset (1876–1937), mais conhecido pelo pseudônimo de Student, para tirar suas dúvidas em relação ao tratamento das pequenas amostras.

Para esse fim, Gosset obteve licença da Cervejaria Guinness, em Dublin, na Irlanda, considerada a maior da época, para fazer um estágio no University College, em Londres, no ano acadêmico 1906/1907, sob a orientação de Pearson. Seu artigo, Student (1908a), é considerado o marco inicial do estudo das pequenas amostras. Nele, Student inicia com o pensamento lapidar de que qualquer experimento pode ser considerado como um indivíduo de uma população de experimentos realizados sob as mesmas condições. Uma série de experimentos é uma amostra extraída dessa população. Esse trabalho surgiu da necessidade de se determinar dentro de que limites a média da população μ estaria situada, conhecida a média da amostra \bar{X} e seu erro-padrão. O método usual era admitir

que $\frac{\bar{X} - \mu}{s/\sqrt{n}}$ obedecia à distribuição normal e utilizar as

tábuas de integrais de probabilidade, mas Student intuiu que, para n pequeno, s estaria sujeito a um erro de amostragem maior e assim isso só seria válido para grandes amostras, pois o valor de s é praticamente equivalente a

σ , sendo portanto $\frac{\bar{X} - \mu}{s/\sqrt{n}} \sim N(0,1)$. Tendo levado em

conta a distribuição de $s^2 = \frac{\sum (X - \bar{X})^2}{n}$ que ele

mostrou, por meio do cálculo dos momentos, ser do Tipo III das curvas de Pearson (em essência uma distribuição do χ^2) e ademais que s^2 e \bar{X} eram independentemente distribuídas, Student derivou a

distribuição de $Z = \frac{\bar{X} - \mu}{s}$, que devido a Fisher passou

mais tarde a ser conhecida por $t = Z\sqrt{n-1}$, sendo

$$t = (\bar{X} - \mu) / \sqrt{\frac{\sum (X - \bar{X})^2}{n(n-1)}}$$

O próprio Student denominou-a t e veio a ser conhecida como distribuição t de Student. Deve ser consultado, a esse respeito, o artigo de Churchill Eisenhart (1979). A contribuição de Student não foi devidamente apreciada na época, pois para Karl Pearson, as pequenas amostras não eram fidedignas, devendo ser evitadas.

Foi Fisher quem reconheceu o mérito desse trabalho, ao qual emprestou seu gênio para desenvolvê-lo teoricamente. Gosset publicou ainda vários trabalhos, sempre com o pseudônimo de Student, imposição da Cervejaria Guinness, que não permitia a seus técnicos usarem seus próprios nomes em suas publicações. Mas isso será assunto para a seção seguinte, A Fase da Experimentação.

No período em que exerceu a cátedra, Pearson dedicou boa parte de seu tempo à construção de tabelas estatísticas, numa época em que predominava o uso das

máquinas de calcular manuais! Foram então publicadas as seguintes tabelas: *Tables for Statisticians and Biometricians* (1914–1931), *Tables for the Incomplete Gama Function* (1922), *Tables for the Incomplete Beta Funcion* (1934), que se revelaram de grande utilidade àqueles que pesquisavam em estatística. Além disso, esteve ocupado com suas funções editoriais da *Biometrika*, atividade que exerceu até sua morte, em 1936. Para deixar essa revista livre somente para artigos de Estatística, K. Pearson fundou outra, *The Annals of Eugenics*, em 1930, dedicando-a exclusivamente à Eugenia e Genética Humana.

Com a aposentadoria de Karl Pearson, seu departamento foi desdobrado em dois: o Departamento de Eugenia e o Departamento de Estatística. O primeiro foi oferecido a R. A. Fisher, que o ocupou como professor catedrático Galton de Eugenia, famoso também nesse assunto, com a publicação do seu livro *The Genetical Theory of Natural Selection*, Fisher (1930b). Nesse cargo, Fisher passou a editor da nova revista *The Annals of Eugenics*, que sob sua influência tornou-se rapidamente importante em artigos sobre estatística. Somente em 1943, quando Fisher saiu para lecionar na Universidade de Cambridge, Inglaterra, seu substituto no Laboratório Galton, L. S. Penrose, dedicou novamente essa revista inteiramente a assuntos de genética humana, modificando seu título para *Annals of Human Genetics*, mantido até hoje.

O Departamento de Estatística ficou a cargo de seu assistente e filho Egon Sharpe Pearson (1895–1980), primeiramente como reader e depois como professor titular. K. Pearson continuou como editor da *Biometrika*, até sua morte, 3 anos depois.

A Fase da Experimentação



Diferentemente das técnicas estatísticas utilizadas pela Escola Biométrica, as pesquisas científicas de natureza experimental exigiam tratamento adequado às pequenas amostras, com objetivo inferencial, conforme revelaram os trabalhos pioneiros de Gosset, que se tornou conhecido pelo seu pseudônimo de Student. Esses trabalhos foram continuados no mais alto nível teórico por R. A. Fisher, a figura mais representativa da Fase da Experimentação, considerado o criador dos métodos modernos da Análise e Delineamento de Experimentos.

Ronald Aylmer Fisher (1890 – 1962) nasceu em Londres, no dia 17 de fevereiro de 1890, e faleceu em Adelaide, na Austrália, a 29 de julho de 1962. Após terminar seus estudos secundários em Harrow, que com Eton eram as mais famosas escolas públicas da Inglaterra (que, na verdade, de públicas só têm o nome), matriculou-se, em 1909, no curso de Matemática da Universidade de Cambridge, com uma bolsa de estudos para o Gonville and Caius College. Em 1912, bacharelou-se, mas continuou por mais 1 ano estudando Mecânica Estatística e Física Quântica com Sir James Jeans, e Teoria dos Erros, com F. M. J. Stratton, ambos renomados astrônomos. Como estudante, interessou-se também pela Genética, tendo lido os trabalhos de Karl Pearson sobre *Mathematical Contributions to the Theory of Evolution*, e participado ativamente como presidente da Sociedade de Eugenia da Universidade de Cambridge.

Durante sua vida profissional, iniciada em 1912, ainda como estudante universitário, e que se estendeu por

meio século, até sua morte, fez contribuições teóricas fundamentais à Estatística e à análise e delineamento de experimentos que, por si, já garantiriam sua fama de um dos maiores cientistas do século. Além dessas contribuições, Fisher foi também um eminente geneticista, exemplo raro da combinação de talento matemático e de cientista natural no mais elevado sentido, o que bem revela a versatilidade de seu gênio. Basta mencionar seu célebre artigo publicado em 1918 e o famoso livro em 1930, anteriormente citados. Com John Burdon Sanderson Haldane (1892 – 1964) e com o geneticista americano Sewall Wright (1889 – 1988) – conhecido dos estatísticos por sua invenção do path coefficient (coeficiente de percurso) na regressão múltipla – Fisher constituiu o triunvirato que governou os primeiros 25 anos de existência da Genética de Populações. Não devem ser omitidas, também, suas importantes pesquisas serológicas, que culminaram com a elucidação do mecanismo hereditário do sistema Rh de grupos sanguíneos.

Considerado por C. Radhakrishna Rao como o fundador da Estatística Moderna, Fisher foi não somente o maior estatístico de sua época, mas para muitos que conheceram sua obra monumental, é ainda o maior estatístico de todos os tempos. Ao longo de sua eminente carreira, recebeu várias honrarias e distinções acadêmicas, entre outras, o grau de Doutor por sua *Alma Mater* em 1926, o título de Fellow of the Royal Society (F. R. S.) em 1929, e o título honorífico de Sir, em 1952.

Para maiores detalhes de sua vida e de sua obra, recomenda-se a leitura da excelente biografia escrita por sua filha Joan Fisher Box (1978), e dos números da revista *Biometrics* 20(2), 1964 (*In Memoriam* Ronald Aylmer Fisher 1890–1962) e *Biometrics* 46(4) 1990, por ocasião

do centenário de seu nascimento. Para completar, com opiniões críticas bem valiosas, é aconselhável a leitura de artigo póstumo do conceituado estatístico de tendência bayesiana, Leonard Jimmie Savage (1917 – 1971), intitulado: *On Rereading R. A. Fisher* (Savage, 1976), considerado por Oscar Kempthorne a mais requintada lição de estatística que ele já tivera.

Fundamentos teóricos da Estatística

As contribuições de Fisher à estatística começaram com o artigo (Fisher, 1912) no qual empregou o método da máxima verossimilhança no ajustamento de curvas de frequência, mas usou o nome de probabilidade inversa, que o levou a corrigir 10 anos mais tarde, no seu clássico trabalho Fisher (1922b). Nessa publicação, com Fisher (1925c), ele resolveu essencialmente o problema da estimação pontual, aplicando o método da máxima verossimilhança.

Ainda recém-formado, Fisher entrou em contato com os trabalhos de Student (1908b), sobre a distribuição t e a subsequente distribuição do coeficiente de correlação em pequenas amostras.

Nesse trabalho, Student usou um processo de simulação, utilizando uma experiência de amostragem semelhante à que fizera para a verificação empírica das distribuições de s^2 e Z de seu trabalho anterior. Entretanto, coube a Fisher, a solução da distribuição exata de r para $\rho \neq 0$ (Fisher, 1915), em cujo trabalho chegou à transformação $r = \tanh(z)$ (sendo z sua própria distribuição), fruto de sua intuição geométrica.

Fisher continuou seus estudos sobre as distribuições amostrais, tendo produzido várias distribuições exatas e desenvolvido os respectivos testes de significância, incluindo-se a demonstração exata da distribuição de Student e sua própria distribuição z , dando sempre preferência pela representação hipergeométrica, que o levou a dar soluções singularmente independentes de simbolismo algébrico. Mostrou, também, a relação entre essas distribuições e a do χ^2 . Para aplicação nos testes de significância, os diversos valores dessas distribuições foram publicados nas conhecidas tábuas *Statistical Tables for Biological, Agricultural and Medical Research*, de Fisher & Yates (1938), com a distribuição da razão de variância, conhecida por F , assim denominada em sua homenagem a Snedecor e relacionada à distribuição z de Fisher pela expressão $e^{2z} = F$.

O artigo acima mencionado foi o único publicado em *Biometrika*. Outro artigo de Fisher (1921) sobre o erro provável do coeficiente de correlação saiu somente em 1921, na nova revista internacional *Metron*, editada por Corrado Gini. Nessa revista, foi também publicado seu trabalho (Fisher, 1925a), no qual é dada a demonstração matemática da distribuição t de Student, utilizando a representação da amostra num espaço n -dimensional, que o próprio Student recebera de Fisher e havia remetido a Pearson, na expectativa de que fosse publicado em *Biometrika*.

Segundo Yates & Mather (1963), em sua biografia sobre Fisher, o método da máxima verossimilhança foi, sem dúvida, uma das maiores contribuições de Fisher à metodologia estatística. O trabalho relevante sobre esse assunto está contido em seus artigos de 1912, 1922 e

1925, previamente citados e mais ainda em Fisher (1935a), em cuja discussão A. L. Bowley atribui ao estatístico e economista Francis Ysidro Edgeworth (1845 – 1926), prioridade na estimação pela máxima verossimilhança, corroborada, de certo modo, por Savage (1976), no seu artigo *On Rereading R. A. Fisher*, ao declarar que Edgeworth antecipara importantes idéias sobre a verossimilhança, embora de modo obscuro. Contudo, segundo Edwards (1974), Edgeworth não se antecipou a Fisher, uma vez que ele fora bem explícito ao usar a Probabilidade Inversa. Semelhante ponto de vista é esposado por Rao (1962), ao argumentar contra idêntica alegação de Neyman (1951), em dar preferência a Edgeworth. Entretanto, a verdade é que coube a Fisher, o mérito de ter criado o conceito de verossimilhança, diferente de probabilidade, e desenvolvido o Método de Máxima Verossimilhança, aplicando-o na estimação de parâmetros, inclusive na Genética, onde encontrou campo fértil de aplicação.

Os conceitos de consistência, eficiência e suficiência foram claramente definidos. Os estimadores suficientes são considerados os mais desejáveis, mas nem todos os parâmetros têm estimadores suficientes, mas quando estes existem, o estimador de máxima verossimilhança é um estimador suficiente, conforme mostrou Fisher, ao revelar a superioridade de seu método de estimação em relação ao método dos momentos, anteriormente usado, desenvolvido por Karl Pearson. O método dos mínimos quadrados, de Gauss, é equivalente ao método da máxima verossimilhança, quando a distribuição é normal. No dizer de Fisher, “um estimador suficiente é aquele que contém toda a informação contida na amostra, sendo desnecessário considerar qualquer outro estimador”. Matemática-

mente, isso significa que a função de verossimilhança

$\prod_{i=1}^n f(x_i, \theta)$ pode ser fatorada do seguinte modo:

$$\prod_{i=1}^n f(x_i, \theta) = g(x_1, x_2, \dots, x_n / \hat{\theta}) h(\hat{\theta}, \theta)$$

onde a distribuição condicional de \underline{x} dado $\hat{\theta}$ é independente do parâmetro θ e $h(\hat{\theta}, \theta)$ é a distribuição conjunta de $\hat{\theta}$ e θ , sendo $\hat{\theta}$ o estimador de máxima verossimilhança do parâmetro θ .

Outra contribuição de Fisher (1930a) foi o conceito de probabilidade fiducial. Probabilidade inversa é sinônimo de probabilidade a posteriori, obtida pelo teorema de Bayes. A probabilidade fiducial foi concebida por Fisher, para resolver o problema central da inferência estatística; como fazer afirmativas de probabilidade sobre um parâmetro desconhecido, sem utilizar o teorema de Bayes, cuja aplicação requer o conhecimento das probabilidades a priori. Quando essas não eram conhecidas, admitia-se o axioma que invoca o princípio da indiferença ou da razão insuficiente, que é o de considerar iguais as probabilidades a priori. A aceitação desse axioma incluiria o problema da inferência indutiva dentro do âmbito da teoria matemática de probabilidades, desenvolvida para tratar de problemas dedutivos oriundos dos jogos de azar.

O teorema de Bayes, da autoria do Rev. Thomas Bayes (1702 – 1761), foi publicado postumamente, em 1763, por autorização de seu amigo Richard Price, e

reproduzido em *Biometrika* (1958). Laplace não hesitou em incorporar o teorema de Bayes em sua grande síntese e em utilizá-lo quando as probabilidades a priori não eram conhecidas, o que leva, conforme mostrou Fisher, a contradições matemáticas. Aplicado à estimação de um parâmetro, o teorema de Bayes pode ser expresso da seguinte forma:

$$p(\theta/x) \propto p(\theta) \cdot p(x/\theta)$$

conforme Plackett (1966), onde $p(\theta)$ é a distribuição de probabilidade inicial ou a priori do parâmetro θ , $p(x/\theta)$ é a distribuição de probabilidade das medidas x quando θ é dado ou seja, a verossimilhança dos valores de θ quando x é dado, e $p(\theta/x)$ é a distribuição de probabilidade final ou a posteriori para o observado x , omitindo-se a constante de proporcionalidade.

Em palavras:

Probabilidade a posteriori \propto Probabilidade a priori x Verossimilhança

Este é o chamado princípio da probabilidade inversa. A forma preferida de Fisher é $P(x/\theta) = L(\theta)$, sendo L de likelihood, para enfatizar que não é uma função de distribuição de probabilidade.

No argumento fisheriano, a distribuição fiducial do parâmetro corresponde à distribuição dos valores concebíveis do parâmetro, mas não é uma distribuição de frequência no sentido ordinário, donde o nome de distribuição fiducial, que é uma distribuição hipotética, de fé ou crença. Assim, por exemplo, ninguém imagina que haja uma infinidade de valores de μ normalmente

distribuídos em torno de \bar{x} . Conceitualmente, há um único valor de μ desconhecido que dá origem às observações; quando se considera outro conjunto de observações outro valor de \bar{x} é obtido, que origina uma diferente distribuição fiducial de μ . As afirmações de probabilidade derivadas desse tipo de argumento são chamadas probabilidades fiduciais. Entretanto, esse conceito foi considerado obscuro, sendo hoje rejeitado pela maioria dos estatísticos. Ao invés dele, é geralmente aceita a teoria dos intervalos de confiança, proposta por Neyman, como será vista mais adiante. Nessa teoria, não é mencionada nenhuma distribuição do parâmetro, pois seu valor é constante. Obtém-se um intervalo aleatório, por exemplo $[\bar{x} - ts_{\bar{x}}; \bar{x} + ts_{\bar{x}}]$ no qual é depositado certo grau de confiança de que ele contenha o verdadeiro valor do parâmetro. O grau de confiança é estabelecido a priori, de modo arbitrário, por exemplo, 95% ou 99%. O intervalo é computado em cada amostra obtida e é esse intervalo que tem determinada probabilidade de conter o parâmetro.

O interesse de Fisher pelo problema da inferência o perseguiu em toda a sua existência profissional. Sobre esse assunto publicou o livro *Statistical Methods and Scientific Inference* (Fisher, 1956), no qual expõe suas idéias, mas como todas suas obras, é também de difícil leitura. Nesse livro, ele considera de modo mais explícito o significado da probabilidade na inferência indutiva que, para ser usado de modo apropriado, deverá satisfazer as seguintes condições:

- Existência de um conjunto referencial mensurável bem definido.
- O assunto sobre o qual se declara a probabilidade deve pertencer a esse conjunto.

- Nenhum subconjunto relevante pode ser reconhecido.

Se, a partir da informação disponível, podem ser reconhecidos subconjuntos com diferentes probabilidades, então o subconjunto ao qual o assunto em questão pertence deve ser considerado o conjunto referencial. Nesse caso, nenhuma probabilidade referente ao assunto baseado no inteiro conjunto é correta, conforme explica Yates (1964a).

No dizer de C. R. Rao, Fisher foi, sem dúvida, o “arquiteto da análise multidimensional”. Nesse assunto, seus estudos pioneiros foram uma fonte de inspiração de vários trabalhos. Assim, a distribuição de Wishart, deduzida em 1928, foi uma extensão multidimensional do método geométrico usado por Fisher, para o caso bidimensional. Sua função discriminante, utilizada para resolver problemas taxonômicos é, em muitos aspectos, semelhante às linhas desenvolvidas por Harold Hotelling (1895 – 1973), em 1931, na distribuição T^2 da razão de Student, e por Mahalanobis, na Índia, para a estimação da diferença entre populações pela distância D^2 , um aperfeiçoamento do coeficiente de semelhança racial usado por K. Pearson e seus associados em pesquisas antropométricas de cranio-metria. Pode-se incluir, também, o critério Λ de Samuel Stanley Wilks (1906 – 1964), uma generalização multidimensional do teste F (exceto que a contribuição do resíduo é no numerador).

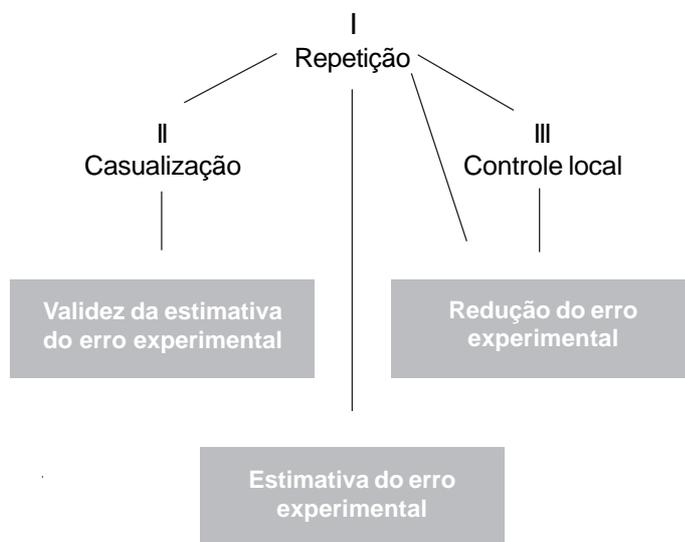
Menção deve ser feita, ainda, às contribuições na análise de correspondência simples e na análise de correspondência múltipla por meio dos seus escores ótimos e das tabelas de contingência bidirecionais com dados categóricos, conforme salienta Gower (1990). Outro

exemplo de teste não-paramétrico é o teste de Fisher de probabilidade exata, aplicado em tabelas de contingência 2 x 2, cuja prática restringe-se às situações em que o número nas células é pequeno, pois, de outro modo, os cálculos tornam-se laboriosos.

Análise de Variância e delineamentos experimentais

A Análise de Variância (Anova) é, provavelmente, o método estatístico de maior repercussão na pesquisa científica, especialmente na experimentação agrícola, de onde surgiu como uma das muitas provas do gênio de Fisher. Em se tratando de uma exposição histórica, é natural que a análise de variância e os delineamentos experimentais, vistos por ele como dois aspectos do mesmo todo, sejam tratados conjuntamente. Seu desenvolvimento e muito de suas aplicações originaram-se no período em que Fisher trabalhou na Estação Experimental de Rothamsted, de 1919 a 1933, a maior e mais antiga das instituições britânicas de pesquisa agrícola onde eram conduzidos ensaios com fertilizantes químicos desde sua fundação, em 1843. A terminologia por ele criada bem reflete essa influência. Suas idéias sobre esse assunto encontram-se em suas duas obras: *Statistical Methods for Research Workers*, (1925b) e *The Design of Experiments*, (1935c), as quais são consideradas suas maiores contribuições à Estatística. Embora destinadas aos pesquisadores das áreas biológica e agrônômica, não são de fácil leitura. Entretanto, graças à sua disseminação foram devidamente interpretadas, especialmente por George Waddel Snedecor (1881 – 1974), autor do livro (*Snedecor*, 1937), que já vai para a sétima edição, tendo W. G. Cochran como co-autor.

Os princípios essenciais do planejamento de experimentos enunciados por Fisher estão representados no diagrama abaixo, afixado na parede do seu laboratório em Rothamsted:



A novidade introduzida por Fisher foi o princípio da casualização, uma brilhante inspiração, inteiramente sua, segundo Yates (1964b). Segundo ele, a casualização garantiria a validade da estimativa do erro e possibilitaria a aplicação dos testes de significância para se verificar o efeito dos tratamentos. Para ele, pela casualização (por um mecanismo objetivo de sorteio) nenhum tratamento seria continuamente favorecido ou desfavorecido nas sucessivas repetições por alguma fonte estranha de variação. Ela se fazia necessária para que as variações que contribuem para o erro experimental fossem convertidas em variáveis aleatórias.

Ainda de acordo com Fisher, a análise de variância pode ser considerada um método estatístico, mas não um teorema matemático, sendo nada mais que um modo conveniente de arranging the arithmetic, segundo suas palavras. Naturalmente, como outras invenções lógicas, é baseada em teoremas matemáticos, previamente demonstrados. Um tratamento matemático elementar encontra-se em Irwin (1931), e uma prova formal mais completa foi dada por Cochran (1934).

A título de rigor histórico, deve ser mencionado que a idéia de comparar a variação entre grupos com a variação dentro de grupos como teste de homogeneidade foi primeiramente estudada por Wilhelm Lexis (1837 – 1914) na Alemanha, no final do século 19, com relação à amostragem de atributos homógrados. Para isso, Lexis criou um critério de dispersão (o termo variância não era conhecido na época), também chamado razão de Lexis, L , ligado ao χ^2 pela relação $\chi^2 / \nu = L^2$, onde ν é o número de graus de liberdade. Contudo, somente graças aos trabalhos de Fisher, essas idéias tiveram maior avanço e pleno desenvolvimento.

O termo variância foi cunhado por Fisher, em 1918, em artigo já mencionado, no qual mostra que as correlações encontradas entre parentes podem ser explicadas pelo mecanismo da herança mendeliana. Neste trabalho foi apresentada uma decomposição percentual da variância total em suas respectivas causas, mas não pode ser considerada uma análise de variância como essa veio a ser conhecida.

O primeiro artigo sobre esse tópico apareceu num artigo de Fisher, com sua assistente Miss Winifred A. Mackenzie (Fisher & Mackenzie, 1923). Trata-se da

análise estatística de um experimento realizado em Rothamsted em 1922, utilizando um delineamento do tipo split plot (subparcela). De acordo com Cochran (1980), em seu artigo póstumo *Fisher and the Analysis of Variance*, Fisher ainda não havia dominado completamente as normas da análise de variância. Na verdade, sua análise denominada Analysis of Variation e não variance, continha erros, pois foi usada apenas uma única estimativa do erro experimental para todas as comparações. É interessante observar que Fisher empregou primeiramente o modelo multiplicativo como mais apropriado. Nesse mesmo artigo, usou também o modelo aditivo, que continuou preferindo em trabalhos posteriores, provavelmente pela maior facilidade de manipulação. Entretanto, 2 anos depois, em 1925, ao publicar *Statistical Methods for Research Workers*, Fisher havia percebido seu erro e apresentou então uma análise estatística correta no parágrafo 42, Ex. 41. Naquela época, ele tinha completo domínio sobre o assunto. Esse exemplo serve para comprovar, mais uma vez, que, na abertura de novos caminhos, os primeiros passos são oscilantes. As idéias nunca surgem de modo pronto e definitivo, obedecendo antes a um processo de desenvolvimento que se aperfeiçoa de maneira tentativa, com hesitações, pois que o processo criativo é formado de conjeturas da imaginação e não uma dedução lógica dos conhecimentos anteriores. Cabe, aqui, a frase do escritor Arthur Koestler The history of ideas is filled with barren truths and fertile errors. (A história das idéias está cheia de verdades estéreis e erros férteis).

Anteriormente, os delineamentos sistemáticos eram os mais populares entre os agrônomos na experimentação de campo, que acreditavam podiam obter maior precisão desse modo. Um exemplo são os quadrados Knut-Vik,

que imitam o movimento da peça do cavalo no jogo de xadrez, uma espécie de quadrado latino. Abaixo, figura um quadrado Knut-Vik 5 x 5,

<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>
<i>D</i>	<i>E</i>	<i>A</i>	<i>B</i>	<i>C</i>
<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>A</i>
<i>E</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>C</i>	<i>D</i>	<i>E</i>	<i>A</i>	<i>B</i>

Quadrado Knut-Vik

onde se vê que nenhum tratamento aparece mais de uma vez numa diagonal. Esse tipo de arranjo era conhecido na Dinamarca desde 1872, mas é atribuído ao norueguês Knut-Vik.

Deve ser dito, que o próprio Student, ao se envolver em experimentação de campo, principalmente na competição de variedades de cevada em larga escala, por motivos de trabalho na Cervejaria Guinness, deu preferência aos arranjos balanceados sistemáticos, argumentando que a casualização causa um aumento da variabilidade. Isso constituiu ponto de discórdia entre Student e Fisher, embora não tenha causado qualquer inimizade entre ambos, que permaneceram amigos até a morte de Student, em 1937. Independentemente de Fisher, ele chegara à estimativa do erro para a comparação de variedades, utilizando, como sempre, sua originalidade,

que Fisher mostrou ser algebricamente equivalente à análise de variância. Em 1923, em sua correspondência com Gosset, Fisher mostrou, também, a derivação do procedimento usado na análise de variância em blocos casualizados por meio do ajustamento de constantes para blocos e para tratamentos pelo método clássico dos mínimos quadrados.

Embora essa tenha sido a primeira abordagem da análise de variância, Fisher deu preferência à apresentação da análise aritmética da decomposição da soma dos quadrados, cuja simplicidade tornou-a acessível aos pesquisadores menos versados em teoria estatística. Isso representou um ganho prático enorme, desde que não fosse elevado a um ritual cego, nas palavras de Maurice Stevenson Bartlett (1910–2002), conhecido por seu teste de homogeneidade de variâncias (Bartlett, 1965). Entretanto, essa simplicidade de cálculos depende do fato de o experimento ter sido delineado para ser ortogonal, i.e., permitir que os efeitos sejam capazes de uma estimação direta e separada, pois, em caso contrário, tornar-se-á necessário usar o princípio clássico dos mínimos quadrados para se estimarem os parâmetros. As técnicas de estimação de parcelas perdidas (missing plot), iniciadas com Allan & Wishart (1930), nada mais são que recursos para restaurar a ortogonalidade, tornando assim possível a análise de variância, segundo padrão simples.

O primeiro reconhecimento ostensivo de que um modelo linear analisado pelo método dos mínimos quadrados era mais fundamental do que uma análise intuitiva baseada no desdobramento da soma dos quadrados deve-se a Yates (1933). Naquela época, Yates ainda não tinha conhecimento da correspondência entre Fisher e Gosset, acima referida. Não demorou para que esse

método fosse utilizado em delineamentos não balanceados e estendido a problemas de regressão múltipla. É pena que a Anova seja estudada atualmente apenas dentro da teoria dos modelos lineares, pela sedução de sua elegância, completamente desligada de sua origem histórica.

As idéias de Fisher foram finalmente vitoriosas e, a partir de 1925, a casualização foi usada como rotina em todos os planos experimentais realizados em Rothamsted. A análise de variância, com a análise de covariância, também por ele desenvolvida, passaram a constituir o instrumental básico para a interpretação dos resultados dos experimentos controlados.

Em seu artigo, Fisher (1926), considerado o precursor de seu livro *The Design of Experiments*, declara, peremptoriamente, que não se deve levar em conta o aforismo de se perguntar à natureza apenas uma questão de cada vez, pois muitas vezes ela se recusa a responder até que outro fator seja acrescentado, advogando, assim, o uso de experimentos fatoriais e suas vantagens. Contudo, ele percebeu as dificuldades práticas para um grande número de fatores, tendo mostrado que essas dificuldades poderiam ser evitadas incluindo-se num bloco apenas uma parte de todas as combinações possíveis. Assim, cada bloco não seria mais uma repetição completa, sacrificando-se deliberadamente a informação de certas interações, consideradas pouco importantes, confundindo-as com as diferenças entre blocos. Esse artifício técnico foi denominado confundimento (confounding), que pode ser total ou parcial, conforme as interações estejam completamente confundidas ou apenas em parte, permitindo, assim, a recuperação da informação sobre as interações confundidas. No supracitado artigo, Fisher mostra, pela primeira vez, sua preferência por um nível de

significância de 5% revelando que talvez outros preferissem um nível mais rigoroso, como 2% ou 1%. Os experimentos fatoriais e as técnicas de confundimento foram posteriormente desenvolvidos por Frank Yates (1902 – 1994) e estão expostos em sua monografia *The Design and Analysis of Factorial Experiments* (1937), à qual deve ser acrescentada sua extensa lista de trabalhos sobre blocos incompletos, em geral.

Em 1931, Yates foi para Rothamsted, em substituição a John Wishart (1898 – 1956), assistente de Fisher desde 1928, que saíra naquele ano para lecionar na Universidade de Cambridge, onde foi responsável pela formação acadêmica de mais de uma geração de ilustres estatísticos. Um exemplo de sua cooperação com Fisher, naquele período, é a publicação de Fisher & Wishart (1930), que mostra a preocupação de ambos em divulgar os novos métodos ao alcance dos pesquisadores. Yates trouxe consigo sua vasta experiência em aplicações do método dos mínimos quadrados de Gauss em levantamentos geodésicos, na antiga colônia britânica da Costa do Ouro (atual Gana); conhecimentos esses que vieram a se tornar de grande utilidade na sua nova função. Foi ativo e eficiente colaborador de Fisher, substituindo-o na chefia do Departamento de Estatística de Rothamsted, quando este se retirou para assumir a cátedra na Universidade de Londres, em 1933, tendo permanecido nesse posto até sua aposentadoria. É fruto dessa colaboração as conhecidas tábuas estatísticas de Fisher e Yates, publicadas em 1938. A elaboração dessas tábuas contou com a eficiente assistência de Wilfred Leslie Stevens (1911 – 1958), conhecido dos brasileiros por suas atividades profissionais neste País, onde faleceu prematuramente como professor da Universidade de São Paulo, cargo que vinha exercendo desde 1948.

Além desses nomes, deve ser lembrado, também, William Gemmell Cochran (1909 – 1980), assistente de Yates por 5 anos, de 1934 a 1939, ano em que emigrou para os Estados Unidos da América, e foi substituído por David John Finney (1917–) ex-assistente de Fisher no Laboratório Galton. Nos Estados Unidos, Cochran exerceu primeiramente o magistério na Universidade Estadual de Iowa, em Ames, onde cooperou com G. W. Snedecor no estabelecimento do Laboratório de Estatística, que ficou famoso como pioneiro das atividades de ensino e pesquisa, sob a direção de Snedecor. Anos depois, Cochran aceitou convite de Gertrude Mary Cox (1900 – 1978) para participar do corpo docente da Universidade Estadual da Carolina do Norte, em Raleigh, de onde saiu para a Universidade John Hopkins e finalmente para Harvard. É preciosa lembrança de sua colaboração com Gertrude Cox a obra de Cochran & Cox (1950), cujo trabalho havia iniciado quando ambos faziam parte do corpo docente de Iowa, de relevantes serviços aos que se dedicam à estatística experimental. Posteriormente, em 1947, veio para Ames, fazer parte do quadro profissional, também procedente da Inglaterra, onde estudara em Cambridge com Wishart e trabalhara alguns anos com Yates em Rothamsted, Oscar Kempthorne (1919 – 2000), outro estatístico de renome. Em seu livro Kempthorne (1952), ele reconhece nas primeiras páginas do prefácio seu débito a Fisher e Yates, cujas contribuições considera como os fundamentos do assunto.

Uma vez estendida a experimentos mais complexos, além de fornecer as estimativas dos erros e os testes de significância dos vários efeitos, a Anova permitiu estimar as componentes de variância atribuídas às diferentes

classes de efeito. Aliás, em *Statistical Methods for Research Workers*, o leitor é introduzido à Análise de Variância nesse contexto, como alternativa à correlação intra-classe e que, segundo o autor, esse método constituía grande simplificação. Fisher achou que a distribuição do coeficiente de correlação intra-classe era essencialmente equivalente à da razão de variâncias. Ele nunca se preocupou de tratar a correlação intra-classe separadamente da análise de variância, nas sucessivas edições de seu livro. Certamente, a forma da análise de variância apropriada à correlação intra-classe ou a qualquer classificação hierárquica requer ampliação para ser usada nas classificações cruzadas da análise de experimentos.

A Anova difundiu-se rapidamente entre os pesquisadores. Para muitos deles, a estimação das componentes de variância era irrelevante, mas em muitos casos essas estimativas tornavam-se necessárias. Tudo isso era bem conhecido até o fim da 2ª Guerra Mundial. Entretanto, depois desse período surgiu novo conceito introduzido por Churchill Eisenhart (1913 – 1994) em artigo sobre pressuposições em que se baseia a análise de variância, Eisenhart (1947). Nesse artigo, ele distingue o Modelo I ou de efeitos fixos, e o Modelo II ou de efeitos aleatórios, tendo sido depois acrescentado o modelo misto, em que alguns efeitos são fixos e outros aleatórios. A análise estatística é a mesma nos diferentes modelos, mas os testes de significância diferem, de acordo com a expectância dos quadrados médios. Na prática, um modelo é de efeitos fixos, se os tratamentos são deliberadamente escolhidos, ou é de efeitos aleatórios (também chamado de componentes de variância) se é feita uma seleção aleatória dos tratamentos, mas o interesse do pesquisador não se

restringe apenas a eles. Esse último é a forma original da análise de variância.

Fisher fez ainda notáveis contribuições em fascinantes problemas combinatórios relacionados com a enumeração dos quadrados latinos e greco-latinos, onde mais uma vez revelou seu gênio. Algumas histórias interessantes relacionadas com esse trabalho são relatadas por Yates (1975). Segundo Finney, em nenhum lugar a elegância da matemática de Fisher é mais evidente do que nos seus artigos sobre esse assunto, publicados no *Annals of Eugenics* de 1940 a 1945. Grande número das soluções encontradas encontram-se nas Tábuas Estatísticas de Fisher e Yates. O próprio Finney (1945) publicou no *Annals of Eugenics*, seu artigo no qual discute o uso de apenas parte de uma repetição num experimento fatorial, quando o número de fatores é grande. Essa técnica ficou conhecida por *repetição fracionada*, para cujo desenvolvimento contribuiu também Kempthorne (1947). Nesse assunto, é particularmente interessante a leitura do Capítulo V do livro de Fisher *The Design of Experiments* (1935c), na parte referente aos quadrados latinos ortogonais, quadrados greco-latinos e de ordem superior.

Esses resultados estimularam as investigações de um grupo brilhante de matemáticos indianos, sob a liderança de Raj Chandra Bose (1901 – 1987) do Instituto de Estatística de Calcutá, fundado por Prasanta Chandra Mahalanobis (1893 – 1972), em 1931. Conta Bose, que em dezembro de 1937, durante um seminário realizado no Instituto, Fisher, então em visita à Índia, conjecturou que seria possível construir um quadrado hiper-greco-latino para todo o valor de p , que é um número primo ou potência de um primo. Foi objetivo de seu primeiro artigo provar que a suposição de Fisher era correta, o que Bose

conseguiu usando as propriedades dos Corpos de Galois e das Geometrias Projetivas Finitas com eles relacionadas. Em seu artigo Bose (1938) desenvolve um método de construir os quadrados hiper-greco-latinos. Na mesma época W. L. Stevens (1938) publicou seu artigo *The Completely Orthogonalized Latin Square*. Bose continuou esses estudos e fez muitas contribuições sobre o assunto, tendo confirmado a conjectura de Euler de que não existe um quadrado latino 6×6 ortogonal, e mostrou num artigo com Shrikvande (1959), a falsidade da conjectura que generalizava a não-existência de quadrado greco-latino da ordem $4t+2$ para todo $t > 1$. Leonard Euler (1707 – 1783) foi o primeiro matemático a se interessar pelos quadrados latinos, denominados “mágicos” naquela época em que fora publicado seu artigo *Recherches sur une nouvelle espèce de quarrés magiques*, em 1782. Na época daquele artigo, Bose já havia emigrado para os Estados Unidos, onde em 1949, passou a integrar o corpo docente da Universidade da Carolina do Norte, em Chapel Hill. Após sua aposentadoria naquela Universidade, continuou suas atividades de pesquisa e ensino na Universidade Estadual do Colorado até sua morte, em 1987.

Uma das figuras mais eminentes do grupo de matemáticos do Instituto de Calcutá foi Calyampudi Radhakrishna Rao (1920 –), conhecido dos estatísticos por seus inúmeros artigos e pelos livros: *Advanced Statistical Methods in Biometric Research* (1952) e a versão mais teórica dele *Linear Statistical Inference and its Applications* (1973). Além da influência inicial de Mahalanobis em sua carreira, Rao obteve seu Ph.D. em Cambridge, sob a orientação de Fisher e Wishart, influências estas que explicam a escola a que pertence.

São exemplos de sua contribuição no assunto de delineamentos experimentais seus trabalhos em cooperação com K. R. Nair, e seu artigo *General Methods of Analysis for Incomplete Block Designs* (1947), quando era ainda estudante em Cambridge. Rao esteve mais de uma vez no Brasil, sendo a última em 1999, quando recebeu o grau de Doutor Honoris Causa pela Universidade de Brasília, em concorrida solenidade.

Para completar o quadro de diferentes influências estatísticas na Índia, é oportuno mencionar outro grupo liderado principalmente por Pandurang Vasudeo Sukhatme (1911 – 1997), cujos trabalhos tiveram início em 1940, no Indian Council of Agricultural Research – Icar, em Nova Delhi, Índia. Começando como uma seção de Estatística do Icar, em 1959 passou a constituir o Institute of Agricultural Research Statistics, de reconhecida importância por suas funções de pesquisa e atividades de treinamento em nível de pós-graduação. Na esfera internacional, P. V. Sukhatme tornou-se conhecido por suas atividades na Food and Agriculture Organization – FAO –, órgão das Nações Unidas, onde a partir de 1951, exerceu o cargo de diretor da Divisão de Estatística na sede, em Roma, por mais de 20 anos.

As necessidades da experimentação industrial, onde são pesquisados vários fatores representados por variáveis quantitativas, levaram ao desenvolvimento de delineamentos especiais, conhecidos na literatura pelo nome de Delineamentos de Box, em homenagem ao estatístico inglês George Edward Pelham Box (1919–), residente nos Estados Unidos desde 1956, o maior responsável por seu desenvolvimento. Inicialmente, seu trabalho foi aplicado na determinação das condições ótimas em processos químicos, mas pode ser estendido a situações em que a

experimentação é seqüencial e os erros envolvidos são de pequena magnitude. O interesse era o de ajustar uma superfície de resposta descrita aproximadamente por um polinômio do 2º grau e, sendo assim, não haveria razão especial para se usar um arranjo fatorial completo. Os novos delineamentos, tais como os delineamentos compostos e os rotacionais, permitem testar grande número de fatores em poucas unidades experimentais. Essas idéias estão expostas nos seguintes artigos: *Box & Wilson* (1951) e *Box* (1954). Assim, os polinômios ajustados podem ser usados como funções de produção para calcular a combinação ótima dos insumos a serem utilizados.

Teste de hipóteses estatísticas

A teoria clássica do teste de hipóteses foi fruto da colaboração entre dois eminentes estatísticos Jerzy Neyman e Egon Sharpe Pearson, iniciada quando *Neyman* estagiava no University College para onde fora estudar com Karl Pearson no outono de 1925. Esta colaboração é relatada por Pearson (1970) em *The Neyman-Pearson Story* (1926 – 1934).

Jerzy Neyman (1894 – 1981) foi, indubitavelmente, um dos gigantes da Estatística. Nasceu em Bendery, na região da Bessarábia, então parte da Rússia Czarista, que pertenceu posteriormente à Romênia e finalmente à República Socialista da Moldávia, atual Moldava, da ex-União Soviética, de ascendentes poloneses de credo católico, pertencentes à pequena aristocracia rural, cujas terras haviam sido confiscadas na Revolta de 1863. Note-se que a Polônia não existia como estado soberano desde

1795, quando houve sua partilha entre a Áustria, a Prússia e a Rússia. No verão de 1921, Neyman, pela primeira vez foi à Polônia, que havia ressurgido como país independente ao terminar a 1ª Guerra Mundial, já no posto de assistente de Matemática no Instituto de Tecnologia de Kharkov, na Ucrânia, em cuja universidade se graduara em Matemática, segundo relata sua biógrafa Constance Reid (1982), autora do livro *Neyman-from Life*, que iniciou a escrever quando ele ainda vivia. Em 1924, Neyman já havia obtido o grau de Doutor com a tese cujo título em inglês é *On the Application of Probability Theory to Agricultural Experiments*, fora preparada no período em que trabalhou no Instituto Nacional de Agricultura, em Bydgoszcz (anteriormente, Bromberg). Essa tese, cujo original é em polonês, com sumário em alemão, acha-se parcialmente reproduzida por Neyman (1990). Contudo, sua exposição mais conhecida encontra-se no artigo de Neyman (1935).

O interesse de Neyman em se aperfeiçoar em Estatística levou-o a obter uma bolsa de estudos para estudar com Karl Pearson, mas cedo percebeu que o nível teórico ensinado não era o que ele esperava, o que o levou a prosseguir seus estudos em Paris, onde assistiu às aulas de Félix Édouard Justin Émile Borel (1871 – 1956), Henri Léon Lebesgue (1875 – 1941) e de outros matemáticos eminentes como Paul Pierre Lévy (1886 – 1971) e Jacques Salomon Hadamard (1865 – 1963) sobre assuntos de seu maior interesse, relacionados com a teoria de probabilidades, medida e integração, que já o haviam entusiasmado quando estudante do grande probabilista russo Sergey Natanovich Bernstein (1880 – 1968) na Universidade de Kharkov. É oportuno registrar que nessa

época ainda não tinha havido a axiomatização do cálculo de probabilidades, que só se deu em 1933, com a obra do matemático russo Andrey Nikolayevich Kolmogorov (1903 – 1987), intitulada em alemão *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Assim, a colaboração com Egon Pearson, iniciada em 1926, teve de ser prolongada por correspondência entre Londres e Varsóvia e encontros ocasionais na França e na Polônia, onde foram discutidas as bases da Teoria dos Testes de Hipóteses Estatísticas, de onde resultou uma série de artigos, assinados por ambos, a saber: *Neyman & Pearson* (1928, 1933a, 1933b, 1936). Nesses artigos, os autores introduziram várias idéias e conceitos novos nos testes de significância da prova de hipóteses. O trabalho foi grandemente facilitado com a mudança de Neyman para Londres, ao aceitar o posto de assistente de Egon Pearson, em 1933, na ocasião em que este havia assumido a cátedra no University College, com a aposentadoria de seu pai. Contudo, já em 1935, foi promovido ao posto de reader, lugar que ocupou até 1938, quando foi para os Estados Unidos, onde fixou residência.

Na teoria de Neyman e Pearson, no problema do teste de hipóteses estão envolvidas duas hipóteses: a que está sendo provada, denominada hipótese nula ou de nulidade, como alguns preferem (null hypothesis) H_0 e a complementar, a hipótese alternativa H_a . Um bom teste estatístico deve ter uma pequena probabilidade de rejeitar H_0 se esta for verdadeira e uma grande probabilidade de rejeitá-la se ela for falsa, o que conduz naturalmente a dois tipos de erros, o erro do Tipo I ou de 1ª espécie correspondente ao primeiro caso, conhecido por nível de significância α e o do Tipo II ou de 2ª espécie β , que seria o de aceitar H_0 se ela for falsa. O quadro abaixo

sintetiza a natureza dos erros envolvidos no processo de decisão por meio dos testes de significância:

	H_0 verdadeira	H_0 falsa
Rejeição H_0	Erro do tipo I (α)	Decisão correta
Aceitação H_0	Decisão correta	Erro do tipo II (β)

A função $1-\beta$ é chamada potência do teste e corresponde à probabilidade de rejeitar H_0 quando esta é falsa. Um teste estatístico corresponde assim à partição do espaço amostral em duas regiões: a região de rejeição de H_0 (região crítica) R e a região de aceitação de H_0 (A aceitação de H_0 corresponde à rejeição de H_a e vice-versa). Simbolicamente, a probabilidade $P(S \in R/H_0) = \alpha$, sendo S o ponto amostral no espaço n -dimensional constituído pelas observações $\underline{x} = (x_1, \dots, x_n)$. Como não é possível escolher uma região crítica que torne mínimo os dois tipos de erros, o teste ótimo segundo a teoria de Neyman-Pearson é aquele, se existir, que para um determinado nível de significância α é uniformemente mais potente (UMP). Esses autores fizeram uma distinção entre hipótese simples e hipótese composta segundo, respectivamente, especificam todos os parâmetros da distribuição ou não. Deve ser citado aqui o reconhecimento de Pearson (1938) sobre a influência que Gosset exerceu sobre seu trabalho, ao mencionar no seu artigo *Student as a Statistician*, que uma carta que ele havia escrito, em 1926, continha em seu primeiro parágrafo o germe da idéia que formou a base das pesquisas de Neyman e Pearson

sobre testes de hipóteses. Foi a simples sugestão de que a única razão válida para se rejeitar uma hipótese estatística seria que alguma hipótese alternativa explicasse os fatos observados com maior grau de probabilidade.

Como método geral para obter o melhor teste estatístico, Neyman e Pearson utilizaram o critério da razão de verossimilhança

$$\lambda = \frac{L(\hat{\omega})}{L(\hat{\Omega})}$$

onde $L(\hat{\omega})$ é o máximo da verossimilhança de H_0 , pertencente ao subespaço ω do espaço paramétrico Ω e $L(\hat{\Omega})$ é o máximo da verossimilhança correspondente a este espaço. O valor λ ($0 < \lambda < 1$) é uma função apenas das observações e sua distribuição quando H_0 é verdadeira deve ser conhecida para que se determine a região crítica. O critério da razão de verossimilhança foi muito frutífero, permitindo a derivação de todos os testes estatísticos conhecidos a partir de uma única regra. Eles provaram, também, que em grandes amostras a quantidade $-2 \log_e \lambda$ obedece à distribuição do χ^2 . A aplicação dos testes de hipóteses pode ser encarada como o resumo de dois cursos de ações possíveis a serem tomadas, resultantes da aceitação ou da rejeição da hipótese de nulidade, fornecendo, assim, o que Neyman denominou de “uma regra de comportamento indutivo”. Assim, estabeleceu o padrão original que levou Abraham Wald (1950) a elaborar sua teoria geral da decisão em sua obra *Statistical Decision Functions*.

Ao regressar à Inglaterra, para lecionar no University College, Neyman interessou-se novamente pelo problema

de estimação intervalar, assunto que o havia ocupado desde 1930. Foi quando surgiu sua teoria de intervalos de confiança, publicada no apêndice de seu famoso artigo *On the Two Different Aspects of the Representative Method: the Method of Stratified Sampling and the Method of Purposive Selection* (Neyman, 1934). A identidade numérica dos limites fiduciais de Fisher e os limites de confiança sugeriram a Neyman que suas teorias eram essencialmente a mesma, levando-o a considerar seu trabalho como uma extensão e aperfeiçoamento das idéias de Fisher. Alguns autores chegaram a usar as expressões limites fiduciais e limites confidenciais como sinônimas. Posteriormente, Fisher (1935b), declara que Neyman tentou desenvolver o argumento da probabilidade fiducial de um modo que, lamentavelmente, ignorava os resultados da teoria da estimação à luz do que fora originalmente lançada. Assim, no chamado teste de Behrens, estudado por Fisher nesse artigo, a diferença entre as duas teorias é acentuada. O teste de Behrens, primeiramente estudado por W. V. Behrens (1929), consiste no teste da diferença de duas médias, cujas variâncias podem ser diferentes. Esse teste requer o uso de tábuas especiais, construídas por Sukhatme e apresentadas nas Tabelas Estatísticas de Fisher e Yates, anteriormente citadas. Para esse teste, não há solução na teoria de Neyman. Para a compreensão completa dessas diferenças, é aconselhável a leitura do artigo de Neyman (1941), no qual ele elucida, de forma didática, a diferença entre as duas teorias, conceitualmente diversas.

Para Fisher, a teoria de Neyman e Pearson servia bem de modelo para a inspeção de qualidade por amostragem, mas não para a pesquisa científica. A região de aceitação e de rejeição de lotes, com as idéias de riscos

do consumidor e do produtor, têm origem nos dois tipos de erros. O erro do tipo I corresponderia ao risco do produtor, i.e., à probabilidade de que uma partida de boa qualidade venha a ser rejeitada, enquanto o erro do tipo II seria a probabilidade de uma partida de má qualidade ser aceita na inspeção, que é o risco do consumidor. Segundo Fisher, deve-se pensar nos problemas científicos em linguagem diferente da exigida na eficiência tecnológica. Por isso, ele não concordou com Wald, ao considerar o tratamento do delineamento experimental como parte do problema geral da decisão. Para Fisher, a decisão está relacionada com a ação prática, enquanto a inferência com o conhecimento científico. Um exemplo familiar da primeira é o controle estatístico de qualidade industrial. Para ele, os intervalos de confiança de Neyman nada afirmam sobre a probabilidade do parâmetro em relação ao resultado obtido numa amostra particular, mas sobre o intervalo calculado sobre hipotéticas amostras que não foram realmente obtidas. Apesar dessas críticas, a inferência estatística foi conduzida de modo consistente com a teoria de Neyman-Pearson, usando testes estatísticos introduzidos por K. Pearson e Fisher.

Abraham Wald (1902 – 1950) em seu artigo (Wald, 1939), apesar de escrito antes de ele conhecer pormenorizadamente a moderna teoria estatística, já estão presentes as noções mais importantes de sua teoria da decisão. Posteriormente, ele colaborou na formalização da análise e experimentação seqüencial, que vinha sendo usada de modo informal, na inspeção por amostragem no controle de qualidade na indústria. Wald vinha trabalhando nesse assunto desde 1943, condensando os resultados no seu livro *Sequential Analysis*, (Wald, 1947). Nesse livro, é apresentado seu teste seqüencial da razão de

probabilidades, por ele considerado uma extensão das idéias de Neyman e Pearson, achando assim os procedimentos ótimos seqüenciais.

Na análise seqüencial, a decisão para terminar o experimento dependerá, em cada etapa, dos resultados previamente observados. Na verdade, ela é uma parte especial de um problema mais amplo, o da decisão estatística, cuja teoria foi desenvolvida por Wald, na década de 1940, em várias contribuições originais que culminaram em sua obra prima, intitulada *Statistical Decision Functions* (1950), anteriormente citada. Lamentavelmente, o alto grau de sofisticação matemática desse livro limita seu acesso a poucos estatísticos. Uma exposição didática bem acessível é encontrada no Capítulo 7 do livro de Vic Barnett (1973). Em essência, a teoria da decisão constitui o princípio geral para se tomar decisões judiciosas em face da incerteza, incluindo assim todos os problemas que justificam a razão de ser da Estatística. Wald inclui o próprio delineamento experimental como parte dos problemas da decisão, considerando isso como um importante avanço. Entretanto, Fisher (1955) criticou essa idéia, exemplificando que no livro de Wald não são mencionados nenhum dos conceitos básicos sobre delineamentos experimentais, tais como repetição, controle local e casualização, nem citados na bibliografia os trabalhos dos principais autores sobre esse assunto. O próprio livro de Fisher é apenas mencionado, mas nenhum uso é feito dele.

Segundo Wald, uma função de decisão estatística é uma regra de decisão ou estratégia $\delta(x)$, que informa que ação tomar se é observado x . Outro elemento no modelo da teoria da decisão é a função perda (loss function) $l(\theta, \delta(x))$, onde θ é o verdadeiro valor do parâmetro.

O desempenho de δ é medido pela perda média incorrida, denominada função risco $R(\theta, \delta) = E[l(\theta, \delta(x))]$. Note-se que Laplace e Gauss consideraram os erros de observação como perdas e o método dos mínimos quadrados tinha sua justificativa na base de minimizar tais perdas. O problema de seleccionar os melhores procedimentos de decisão tem sido atacado de vários modos. Um deles é pelo critério minimax, ao invés de usar o valor médio do risco. Por esse critério, oriundo da teoria dos jogos desenvolvida de modo determinístico na obra do célebre matemático John von Neumann (1903 – 1957) e do economista Oskar Morgenstern (1902 – 1977) (Neumann & Morgenstern (1944)), é minimizado o risco máximo, daí o nome minimax. Esse critério tem por fim maximizar a proteção contra o pior que possa acontecer, sendo por isso considerado muito conservador. Além da vantagem de conservador, o critério minimax é independente da distribuição a priori do parâmetro e tem risco constante, sob certas condições. Wald esteve à procura de outros critérios, mas sem resultado satisfatório. Ele deu ainda muitas contribuições à Estatística, mas infelizmente faleceu aos 48 anos num acidente aéreo na Índia, de onde regressava após lecionar sobre os temas de seu recém-publicado livro. Esta fatalidade privou a comunidade científica de um possante intelecto, no auge de sua carreira, tendo muito ainda para contribuir para o desenvolvimento da estatística matemática.

Por iniciativa de William Edwards Deming (1900-1993), o Departamento de Agricultura dos Estados Unidos convidou Neyman a fazer uma série de palestras na Escola de Pós-Graduação em Washington, no verão de 1937. Foi nesta oportunidade que ele aceitou o convite para organizar e dirigir um Laboratório de Estatística no Departamento de Matemática da Universidade da

Califórnia, no campus de Berkeley, na posição de professor titular (full professor). Essa nova atividade teve início em agosto de 1938, tendo continuado ao longo dos anos, mesmo depois da aposentadoria, em 1961, como professor emérito, até sua morte em 1981. Já em 1955, fora criado o Departamento de Estatística ao qual ficou agregado o mencionado laboratório. Nesse período, Neyman conseguiu fazer o que veio a ser considerado o maior e melhor centro mundial de ensino e pesquisa da estatística, tendo atraído para Berkeley, sob sua liderança, uma plêiade de profissionais da mais alta categoria, tais como: Erich Leo Lehmann (1917 –), autor do livro *Testing Statistical Hypotheses* (1959) e, mais recentemente, *Theory of Point Estimation* (1983), Lucien Le Cam (1924 –), Henry Scheffé (1907 – 1977), conhecido pelo teste estatístico que tem seu nome e pelo livro *The Analysis of Variance* (1959), David Blackwell (1919 –), Michel Loève (1907 – 1979), conhecido probabilista, e muitos outros. Ficaram famosos os simpósios promovidos quinzenalmente, a partir de 1945, nos quais participaram os mais ilustres colaboradores, cujos trabalhos foram publicados posteriormente, anualmente, como *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*.

Nessa fase, a última de sua carreira profissional, Neyman continuou ocupado com suas pesquisas anteriores relacionadas com a inferência estatística, como atestam suas publicações sobre os melhores estimadores assintoticamente normais (best asymptotically normal – BAN) e sobre os testes ótimos de hipóteses compostas, os chamados C -alfa testes, sendo C em homenagem a Harald Cramér (1893 – 1985). Entretanto, sua maior produção foi em trabalhos aplicados nas áreas de astronomia, meteorologia e medicina, merecendo destaque

seus artigos com Elizabeth Leonard Scott (1917 – 1988) referentes à distribuição espacial de galáxias e do enfoque estatístico aos problemas da cosmologia, e os seus próprios sobre análise estatística dos resultados dos ensaios para provocar chuvas artificiais que deram origem ao artigo de Neyman (1967). Na última área, os melhores trabalhos foram condensados em *Probability Models and Cancer* de Neyman & Le Cam (1982).

A influência de Neyman em Berkeley, ao lado da de Wald, na Universidade de Columbia, foram responsáveis pela maior formalização matemática da estatística nos Estados Unidos, dando assim início à separação da origem inglesa, de predominância fisheriana. Ele fez também estudos pioneiros sobre levantamentos por amostragens, onde introduziu novos métodos, mas esse assunto será visto na próxima seção.

Desenvolvimento dos Levantamentos por Amostragem



esta seção, será tratada a coleta de informações na forma como se apresentam, sem controle das fontes de variação por parte do pesquisador, i.e., dados de natureza não experimental. Cronologicamente, tiveram origem de certo modo desde os primórdios da Estatística, mas na verdade só se desenvolveram após as contribuições provenientes da teoria da estatística, que se desenvolveu principalmente na fase de experimentação, vista na seção anterior. Vale salientar que enquanto nos experimentos lidam-se com populações imaginárias e infinitas, nos levantamentos as populações são reais e finitas.

O estudo das populações finitas é bem mais recente, a teoria é mais difícil e as fórmulas são mais complicadas. A esse respeito, é interessante a leitura do artigo de Hansen & Hurvitz (1943). As necessidades práticas da amostragem em levantamentos incentivaram estudos teóricos dos quais surgiu uma grande variedade de planos de amostragem, que não teriam ocorrido se a questão da amostragem tivesse ficado restrita unicamente às populações infinitas. Os levantamentos por amostragem são praticamente o único método de pesquisa nas ciências econômicas e sociais. Um relato histórico das aplicações da amostragem em levantamentos é encontrado nos seguintes artigos: Stephan (1949), Seng (1951) e Hansen & Madow (1976). Uma visão geral resumida sobre o assunto pode ser vista no Capítulo IV *Les Sondages*, do livro de Dreesbeke & Tassi (1990).

Quem primeiro advogou o uso da amostragem em levantamentos (sample surveys) foi Kiaer, com seu método de representatividade, na reunião do Instituto Internacional de Estatística (criado 10 anos antes, em Londres) em 1895, em Berna, Suíça. Andres Nicolas Kiaer (1838 – 1919) era então Diretor do Bureau Central de Estatística em Cristiania, como era chamada a capital da Noruega, que em 1925 passou a se chamar Oslo. A idéia de Kiaer sobre amostra representativa era a de ser uma miniatura aproximada da população. O método que ele propôs correspondia ao que, atualmente, seria um bem trabalhado método de estratificação, levando-se em conta fatores geográficos, sociais e econômicos. Além disso, ele introduziu uma seleção proporcional em cada estrato, baseada em detalhes do prévio censo demográfico. Kiaer defendeu sua idéia nas várias reuniões do Instituto Internacional de Estatística, realizadas em São Petersburgo,

Rússia, em 1897, em Budapeste, Hungria, em 1901, e em Berlim, Alemanha, em 1903.

O ponto final da polêmica sobre representatividade pode ser situado na Reunião de Roma, em 1925, na apresentação do relatório de Adolphe Jensen, Chefe do Departamento de Estatística da Dinamarca, sobre as conclusões da Comissão ad hoc nomeada no ano anterior, composta pelo próprio Jensen e os membros Arthur Bowley, Corrado Gini, Lucien Mach, Verrijn Stuart e Franz *Žižek*, para estudar o Método Representativo. Em resumo, o período de 1895 a 1925 marcou uma transição nas estatísticas governamentais de plena confiança na cobertura total fornecida pelos censos para a obtenção de informações adicionais nas investigações parciais. Entretanto, reinava ainda confusão sobre o significado da amostragem representativa. Uma descrição detalhada desses fatos encontra-se no artigo de Kruskal & Mosteller (1980).

Deve-se, principalmente, a Arthur Lyon Bowley (1869 – 1957), professor de Estatística na London School of Economics, a teoria e a possibilidade de aplicações práticas do método representativo. Bowley considerou a amostra representativa como podendo se realizar por dois métodos: seleção intencional (*purposive*) ou ao acaso (amostra aleatória), podendo haver combinação dos dois. No caso da amostra intencional, ela se forma mediante uma seleção de grupos de unidades que se supõe dão à amostra as mesmas características que tem a população. Esse método era aconselhado quando não era possível a seleção ao acaso, sendo às vezes mais econômico, mas com a desvantagem de não poder ser calculado o erro de amostragem. Outra aplicação da amostragem

representativa foi feita pelos estatísticos italianos Corrado Gini (1884 – 1965) e seu assistente Luigi Galvani, para obterem uma amostra dos dados do censo de 1921 em 1926/1927. Na época, Gini era o maior estatístico italiano, editor da revista *Metron*, internacionalmente reconhecida, e de grande influência na direção do Escritório Central de Estatística de seu país. Eles decidiram retirar os dados de 29 das 214 unidades administrativas (circondari) em que a Itália estava dividida. Essas 29 unidades eram tais que seus valores médios de sete importantes características eram próximos das médias do país inteiro. Entretanto, quando outras características foram consideradas, ou quando outros aspectos que não apenas as médias, por exemplo a variabilidade ou as associações das sete características foram levados em conta, Gini e Galvani encontraram grandes discrepâncias entre os dados da amostra e do país como um todo. Gini (1928) concluiu em seu artigo sobre tal método de amostragem que os controles eram ineficientes, crítica idênticamente repetida no artigo de Gini & Galvani (1929).

Essa era a situação quando foi publicado o famoso artigo de Neyman (1934), citado na sessão anterior. Nesse trabalho, considerado uma contribuição pioneira, um verdadeiro divisor de águas, no dizer de Kruskal e Mosteller no artigo anteriormente citado, Neyman mostrou a superioridade da amostragem aleatória estratificada sobre a seleção intencional, então aconselhada como método representativo da amostragem e severamente criticada por Gini e Galvani, cuja aplicação no censo italiano é citada, detalhadamente, em seu artigo. A descrença no método representativo, argumentou Neyman, devia ser na amostragem intencional e não na amostragem probabilística, que ele recomendava como a do método

representativo. Na verdade, não há processo seguro para se saber se determinada amostra é representativa de uma população. Operacionalmente, define-se como representativa uma amostra selecionada ao acaso, de modo que cada unidade de amostragem tenha uma probabilidade conhecida, e diferente de zero, de participar na amostra.

O artigo de Neyman é repleto de boas idéias. Nele, é apresentada uma discussão sobre a inferência em populações finitas na base da casualização introduzida pelos procedimentos de seleção. Pela primeira vez, em inglês, é tratada a estimação intervalar pelos intervalos de confiança, conforme já foi visto na seção anterior. É pena que a discussão sobre esse assunto, tratado no apêndice do artigo, tivesse tirado de certo modo o brilho da apresentação da parte principal do trabalho sobre métodos de amostragem. Nesta sessão, que contou com a presença de Bowley, Fisher e E. S. Pearson, entre outros membros da sociedade, houve apenas pequenas discordâncias salientadas por Pearson que os dois métodos não eram exatamente iguais, e reforçadas por Fisher, ao declarar que o argumento fiducial era utilizado somente nos casos de existir uma estatística suficiente, publicadas posteriormente em seu artigo sobre o assunto, conforme foi citado na seção anterior.

Foi na apresentação do trabalho de Neyman (1935), sobre *Statistical Problems in Agricultural Experimentation*, com a colaboração de K. Iwaskiewicz e St. Kolodzieczyk, apresentado em memorável sessão em 28 de março de 1935, que teve início a desavença com Fisher. Essa e outras discordâncias resultaram em polêmicas entre Fisher e Neyman que culminaram, lamentavelmente, com o desentendimento entre ambos. O procedimento adotado nessas apresentações era o de circular o manuscrito com

antecedência entre os membros convidados para participar na discussão, para que pudessem preparar seus comentários na discussão que se seguia após a apresentação formal do trabalho. As discussões, juntamente com a resposta do autor do trabalho, eram finalmente incorporadas ao texto do artigo publicado, tornando possível sua reprodução em futuras citações dos historiadores da ciência para elucidar eventuais dúvidas.

Neyman tratou, também, da amostragem estratificada, que antes havia sido considerada por Bowley, tendo discutido e deduzido a partilha ótima (optimum allocation) das unidades da amostra nos diferentes estratos. Esse assunto fora antes estudado pelo russo Alexander Alexandrovitch Tchuprov (1874 – 1926) no seu artigo (1923), mas não era do conhecimento de Neyman, nem teve repercussão prática nos levantamentos por amostragem. Contudo, Neyman (1952), reconheceu publicamente a prioridade de Tchuprov tão cedo tomou conhecimento. Nessa área, outra importante contribuição de Neyman foi seu artigo sobre amostragem dupla, ou em duas fases, Neyman (1938) cujo título é: *Contributions to the Theory of Sampling Human Populations*. Esse artigo, ao lado de seu famoso artigo acima mencionado, seriam suficientes para garantir a Neyman o reconhecimento de sua preeminência na teoria dos levantamentos por amostragem.

A noção intuitiva de miniatura da população deu origem a processos de escolha das amostras por conveniência ou de amostras intencionais, por apreciação subjetiva. Um exemplo do tipo mencionado é a amostragem por quotas, que é um método de amostragem estratificada, em que a seleção dentro dos estratos não é ao acaso. Essas quotas (estratos) são em número tal que sua

proporção na amostra é aproximadamente a mesma que na população. O argumento contrário ao seu uso é que não se pode calcular o erro de amostragem. É aconselhável a leitura do artigo: *An experimental study of quota sampling* de Moser & Stuart (1953). Contudo, a amostragem por quotas foi largamente utilizada em pesquisas de mercado e de opinião pública, como nas intenções de voto em pesquisas eleitorais. A prevista vitória de Thomas Dewey na disputa com Harry Truman nas eleições de 1948, para a presidência dos Estados Unidos, descreditou publicamente o método de amostragem por quotas, usado por George Gallup, jornalista a cargo da pesquisa. Apresentadas as causas do fracasso, Gallup abandonou o método de amostragem por quotas, de caráter não probabilístico, e passou a utilizar um plano de amostragem onde em todas as suas etapas prevalecia o conceito de aleatorização, conforme escreve Jorge de Souza (1990), professor titular de Estatística da Universidade de Brasília aposentado, em sua obra *Pesquisa Eleitoral Críticas e Técnicas*, cuja leitura é recomendada aos interessados neste assunto.

Na prática, pelo menos em levantamentos sociais e econômicos, raramente uma amostra é selecionada de modo estritamente aleatório. Geralmente dispõe-se de uma lista de unidades de amostragem e usa-se algum tipo de seleção, como por exemplo, cada décima unidade da lista ou outra forma de seleção sistemática.

A amostra obtida pelo emprego de um método de seleção que consiste em sortear apenas a primeira unidade, sendo as demais selecionadas segundo um plano sistemático, é conhecida, também, pela denominação de amostra quase-aleatória (quasi-random sample). Esse termo é citado em Buckland (1951) e em Moser (1961),

na pág. 76, sendo também mencionado por Yates (1946). Um estudo teórico da amostragem sistemática encontra-se em: Madow & Madow (1944).

Os levantamentos por amostragem tiveram posteriormente desenvolvimento em diversos países, principalmente nos Estados Unidos, nas atividades exercidas pelo Bureau of the Census, órgão de longa história, com trabalho pioneiro no desenvolvimento, construção e aplicação do equipamento de processamento de dados em cartões perfurados, conhecidos como cartões Hollerith, nome de seu inventor Herman Hollerith (1860 – 1929). Atualmente, esse processamento é feito por computação eletrônica. Uma relação das atividades do Bureau of the Census encontra-se no artigo de Morris Howard Hansen (1910 – 1990) *Some History and Reminiscences on Survey Sampling* (1987), que participou por vários anos do quadro profissional do Bureau. Sua experiência, com a de William Hurwitz e William Madow, é refletida no livro em dois volumes de Hansen et al., (1953). O último desses autores lecionou na Universidade de São Paulo, em 1946 e 1947, tendo ministrado durante o período de férias acadêmicas, de dezembro de 1946 a fevereiro de 1947, um curso intensivo sobre a Teoria dos Levantamentos por Amostragem, no Rio de Janeiro, a técnicos brasileiros de estatística, a convite do IBGE. As notas de aula deram origem ao livro *Teoria dos Levantamentos por Amostragem* (Madow, 1951). Ao lado da edição brasileira, foi publicada pelo Instituto Nacional da Estatística de Portugal, a edição portuguesa da mesma obra.

Embora muitas idéias usadas na teoria dos levantamentos por amostragem sejam oriundas dos trabalhos de Fisher, como casualização e controle local

(estratificação), ele pessoalmente não escreveu qualquer livro sobre técnicas de amostragem, especialmente nas aplicações em levantamentos. Entretanto, enquanto estava em Rothamsted, Fisher estudou o uso da amostragem em parcelas experimentais, com conseqüências no desenvolvimento e melhoramento das estimativas de produção agrícola e das áreas cultivadas. Depois da Segunda Guerra Mundial, Fisher foi membro da Subcomissão das Nações Unidas sobre Amostragem Estatística onde, ao lado de Mahalanobis e Yates, teve grande influência nos trabalhos publicados pela Subcomissão, especialmente na terminologia recém-criada. O próprio livro de Yates, *Sampling Methods for Censuses and Surveys* (1949), surgiu da solicitação da Subcomissão na sua primeira sessão, em Lake Success, em Long Island, New York, em setembro de 1947, para que fosse preparado um manual para auxiliar o projetado Censo Mundial da População e da Agricultura, em 1950. Esse livro, o primeiro sobre o assunto, e cuja quarta e última edição foi publicada em 1981, preencheu plenamente as necessidades previstas pela Subcomissão das Nações Unidas e é recomendável ser consultado, especialmente em problemas práticos relacionados com o planejamento e a execução dos levantamentos por amostragem.

Na Grã-Bretanha, a necessidade de obtenção de dados por meio de levantamentos não foi tão urgente como nos Estados Unidos. A população do país é relativamente menor e muito mais concentrada. Muitos dados sobre a população são obtidos como subproduto da administração, não necessitando de levantamentos especiais. O primeiro uso de técnicas de amostragem no país foi feito por Bowley, num levantamento em Reading,

em 1912, no qual ele tomou aproximadamente cada 20º domicílio da classe operária, com atenção no cálculo dos erros de amostragem e na possibilidade de viés (bias) pela introdução de substituições e recusas, conforme publicado em seu artigo (Bowley, 1913). Contudo, na Inglaterra, a amostragem nunca alcançou a aceitação que teve nos Estados Unidos, segundo Moser (1949, 1955), cujos artigos tratam, minuciosamente, desse assunto.

A Índia é exemplo de outro país que contribuiu para o desenvolvimento e aplicação dos levantamentos por amostragem com finalidade especialmente para o planejamento econômico na década de 1930, sob a orientação do Instituto Indiano de Estatística organizado por P. C. Mahalanobis. Essas atividades deram origem a seus artigos *Mahalanobis* (1944, 1946). Entre outras atividades, Mahalanobis introduziu a técnica de sub-amostras interpenetrantes para controlar e avaliar a contribuição de erros alheios à amostragem, didaticamente explicada no livro texto de Cochran (1953), no Capítulo 13 *Sources of Error in Surveys*. Uma lista de erros alheios à amostragem é citada por Deming (1950), e algumas medidas para controlar esse erro em levantamentos realizados em populações humanas são estudadas por Hansen & Steinberg (1956). Note-se que o erro de amostragem é apenas uma pequena porção do erro total.

No Indian Council of Agricultural Research – Icar –, P. V. Sukhatme fez importantes trabalhos em levantamentos agrícolas, para estimar a produção numa série de crop-cutting surveys. Foi quando surgiram controvérsias entre ele e Mahalanobis sobre o tamanho das parcelas. Para Sukhatme, as parcelas pequenas, usadas como unidade de amostragem, eram sujeitas a vieses (biases), devido à tendência de serem nelas incluídas as

plantas limítrofes da parcela, havendo assim uma superestimativa da produção inversamente proporcional ao tamanho da parcela, conforme mostrou em Sukhatme (1946, 1947). Posteriormente, Sukhatme foi para a Organização das Nações Unidas para Agricultura e Alimentação – FAO, em Roma, onde, como diretor da divisão de Estatística, continuou a estender os métodos e a teoria da amostragem e erros nos levantamentos, incorporando a experiência dessa instituição nos países em desenvolvimento na promoção dos censos agropecuários mundiais, na segunda edição de seu livro *Sampling Theory of Surveys with Applications* (Sukhatme & Sukhatme, 1970). Essa nova edição contou com a colaboração de seu irmão B. V. Sukhatme como co-autor, e contém vários exemplos de erros alheios à amostragem.

Já em 1938, nos Estados Unidos, o Departamento de Agricultura e o Laboratório de Estatística da Universidade Estadual de Iowa estabeleceram um programa cooperativo de pesquisa sobre amostragem, dirigido por Arnold J. King e Raymond J. Jessen, que estimulou consideravelmente o desenvolvimento de levantamentos agrícolas. Desses estudos merecem ser citados a clássica publicação de Jessen (1942) e o artigo de King & Jessen (1945) *The Master Sample of Agriculture*. Nesse artigo, é estudado um método que veio a ser conhecido como amostragem por área, a ser usado em levantamentos onde as unidades de amostragem são visitadas, pessoalmente, pelos recenseadores. O Bureau of the Census interessou-se em utilizá-lo em conexão com o Censo Agrícola de 1945, quando ele foi utilizado em larga escala. A extensão desse método às áreas urbanas foi feita em seguida. O Bureau of the Census já havia considerado o uso da amostragem como parte do

censo demográfico decenal, o que teve início no Censo de 1940, com a finalidade de coletar informações suplementares a custo razoável. No Brasil, a amostragem começou a ser usada a partir do Censo Demográfico de 1950, na avaliação das tabulações avançadas, que se fizeram necessárias, dada a demora da publicação dos resultados obtidos no censo completo.

Devem ser mencionadas, também, as atividades do Survey Research Center, que depois se tornou o Institute for Social Research, na Universidade de Michigan, Estados Unidos. Aí, foram treinados no Summer Program for Foreign Statisticians, de 1961 a 1981, mais de 400 estudantes provenientes de 94 países, sob a orientação de Leslie Kish (1910 – 2000), conforme ele próprio declara na entrevista dada a Frankel & King (1996), *A Conversation with Leslie Kish*. Kish publicou vários trabalhos, entre eles um dos primeiros com seu chefe Roe Goodman (Goodman & Kish, 1950). Ele executou levantamentos por amostragem em vários países, principalmente na América do Sul, por sua fluência em espanhol, língua que aprendera durante sua participação na Guerra Civil Espanhola. O livro de Kish (1965) *Survey Sampling* foi editado também em espanhol e traduzido até em chinês, segundo consta.

O Bureau of the Census estendeu suas atividades nos levantamentos por amostragem em várias áreas, aumentando para isso seu quadro profissional de modo considerável, tornando-se mundialmente a instituição mais forte nessas atividades. Como consequência, o *Labor Force Survey* passou a ser efetuado em base de amostragem probabilística com o nome de *Current Population Survey* (Levantamento Periódico da População), incluindo maior número de informações sobre a população.

Com o tempo, o *Levantamento Periódico da População*, passou a servir a várias necessidades, tornando-se um modelo para o planejamento de levantamentos por amostragem a ser usado no mundo inteiro, ao qual foram incorporados os resultados de vários trabalhos teóricos sobre o assunto, tais como: amostragem estratificada multietápica de conglomerados, probabilidades de seleção proporcionais ao tamanho com reposição e sem reposição, uso de informação auxiliar nos estimadores, controle adequado dos erros alheios à amostragem, e dos erros de amostragem. Entre esses trabalhos, destacam-se os de Horvitz & Thompson (1952), Yates & Grundy (1953), Sampford (1962) e outros, citados por Tore Dalenius (1962).

No Brasil, um exemplo dessas atividades é a Pesquisa Nacional por Amostra de Domicílios – Pnad –, implantada progressivamente a partir de 1967, para a obtenção de informações básicas necessárias para o estudo do desenvolvimento socioeconômico do País. A Pnad teve início no segundo trimestre de 1967, sendo os resultados apresentados com periodicidade trimestral até o primeiro trimestre de 1970. A partir de 1971, os levantamentos passaram a ser anuais, com realização no último trimestre. A pesquisa foi interrompida para a realização dos censos demográficos de 1970, 1980, 1991 e 2000. Na década de 70, os principais temas investigados, além das características gerais da população, da educação, do trabalho, do rendimento e da habitação, foram migração e fecundidade. Em 1974/1975, foi efetuada uma pesquisa especial denominada Estudo Nacional da Despesa Familiar, que, além dos temas anteriores, investigou dados sobre consumo alimentar e orçamentos familiares.

A Pnad é realizada por meio de uma amostra probabilística de domicílios, obtida em três etapas de seleção: unidades primárias – municípios; unidades secundárias – setores censitários; e unidades terciárias – unidades domiciliares. Na primeira etapa, os municípios são classificados em dois conjuntos. No primeiro, são selecionados os municípios que, em decorrência do tamanho da sua população ou de alguma característica e importância, participam necessariamente da amostra. No segundo conjunto os municípios passam por um processo de estratificação e, em cada estrato, são selecionados com reposição e com probabilidade proporcional à população residente, obtida no censo demográfico mais recente. Na segunda etapa, os setores censitários são selecionados em cada município da amostra, também com probabilidade proporcional ao tamanho e com reposição, sendo o número de unidades domiciliares existentes por ocasião do último censo demográfico usadas como medida do tamanho. Finalmente, na última etapa, os domicílios são selecionados com equi-probabilidade, em cada setor censitário da amostra, para investigação das características dos moradores e da habitação. Maiores detalhes em relação à Pnad podem ser obtidos nas publicações do IBGE, sobre o assunto.

A Era Atual



A era atual caracteriza-se pelo aumento gradativo de matematização da estatística e da influência crescente do uso dos computadores. Na década de 40, a estatística teórica podia ser compreendida por alguém com

conhecimento razoavelmente, bom em Matemática. A situação mudou muito, a ponto de a maioria dos estatísticos não conseguir, atualmente, ler os artigos publicados, dado seu alto grau de sofisticação matemática, mesmo nas revistas supostamente de caráter aplicado. Entretanto, deve ser dito que a Estatística não é propriamente Matemática, nem mesmo matemática aplicada. Como lida com a coleta, a análise e a interpretação de dados, inclui, naturalmente, muita conjectura sagaz, diferente do rigor da demonstração matemática, para não mencionar o raciocínio indutivo envolvido na inferência estatística. Evidentemente, saber Matemática é importante para um estatístico e quanto mais melhor, pois a teoria estatística não envolve apenas conceitos, necessitando também ser formalizada. Contudo, conhecer Matemática, embora necessário, não é suficiente para formar um estatístico.

Segundo Cox (1997), em seu artigo *The Current Position of Statistics: A Personal View*, os anos de 1925 a 1960 podem se considerados a época áurea do pensamento estatístico. Este período abrangeu a maior parte dos trabalhos sobre inferência de Fisher, Neyman, Egon Pearson e Wald, além do desenvolvimento dos delineamentos experimentais e levantamentos por amostragem, assim como as idéias fundamentais sobre séries temporais e análise multidimensional, e as contribuições bayesianas objetivas de Sir Harold Jeffreys (1891 – 1989) e as subjetivas de Bruno de Finetti (1906 – 1985) e L. J. Savage. O controle estatístico da qualidade e os ensaios clínicos casualizados também já estavam firmemente estabelecidos. Embora tenham sido publicados importantes trabalhos entre 1960 e 1985, esse período foi primariamente de consolidação das idéias anteriormente

desenvolvidas. No início deste período, a maioria dos estatísticos já tinha acesso aos computadores eletrônicos, mas a obtenção dos resultados era ainda tarefa demorada. O expressivo aumento subsequente dos recursos de computação e de sua disponibilidade proporcionaram novos desenvolvimentos e facilitaram sobretudo a implementação dos métodos correntes. Visto num prazo mais longo, houve uma verdadeira explosão do assunto, como mostra a quantidade de trabalhos publicados, o surgimento de novas revistas e a quantidade de profissionais comprometidos na área. Se os estatísticos como um todo continuarem envolvidos em importantes atividades científicas, tecnológicas e de negócios públicos, se novas idéias forem encorajadas e, especialmente, se as alarmantes tendências de fragmentação do assunto puderem ser evitadas, conforme declara Cox, são fortes as perspectivas de um novo período de grandes inovações.

David Roxbee Cox (1924 –) é um dos mais prolíferos estatísticos da era atual, conforme mostra o artigo *A Conversation with Sir David Cox*, de Nancy Reid (1994), co-autora de seu último livro *The Theory of the Design of Experiments* (Cox & Reid, 2000). Uma versão não-matemática desse assunto já havia saído intitulada *Planning of Experiments* (Cox, 1958c), muito apreciada pelos estudiosos das ciências experimentais. Ele é autor e coautor de mais de 200 artigos e mais de 15 livros. Além dos mencionados, merecem ser citados: Cox & Miller (1965), cujo assunto, como é sabido foi iniciado por Andrey Andreyvich Markov (1856–1922), nas chamadas cadeias de Markov; Cox (1970), Cox & Hinkley (1974) e Cox & Oakes (1980), entre outros. Dos artigos publicados, salientam-se os de Cox (1958a; 1958b; 1972), para citar apenas os de maior repercussão.

David Cox graduou-se em Matemática na Universidade de Cambridge, Inglaterra, e obteve seu Ph.D na Universidade de Leeds, em 1949, quando trabalhava na Wool Industries Research Association, em Leeds, também na Inglaterra. De 1950 a 1955, foi assistente em Cambridge, no Statistical Laboratory, dirigido por Wishart, e em seguida lecionou por 15 meses na Universidade da Carolina do Norte, em Chapel Hill, Estados Unidos. Ao regressar à Inglaterra, foi professor adjunto de Estatística no Birkbeck College e, a partir de 1966 até 1988, foi professor titular no Imperial College, ambos na Universidade de Londres. Em 1988 mudou-se para Oxford, onde foi diretor do Nuffield College, posto no qual se aposentou em 1994, sendo, atualmente, membro honorário do mesmo e professor emérito de Estatística na Universidade de Oxford. Em 1973, foi eleito Fellow da Royal Society (F. R. S.) e em 1985, recebeu o título honorífico de Sir. Entre suas honrarias, Sir David recebeu mais de dez doutorados, sendo o último de Doutor Honoris Causa outorgado pela Universidade Federal do Rio de Janeiro, em 28 de julho de 2000. Finalmente, menção deve ser feita ao seu exercício de editor da *Biometrika* de 1966 a 1991, cargo que desempenhou com sua costumeira dedicação e eficiência por 25 anos.

Em seu artigo *Computers – The Second Revolution in Statistics*, Yates (1966) revela que, para ele, a primeira revolução na Estatística veio com a introdução das máquinas de calcular. De fato, tanto as contribuições de Karl Pearson como as de R. A. Fisher, no desenvolvimento teórico da Estatística, não teriam ocorrido não fosse o precioso auxílio prestado pelas máquinas de calcular. Ambos certamente esposam esse reconhecimento, como mostram as fotos por eles tiradas

ao lado de suas calculadoras. Nas décadas de 40 e de 50, as máquinas de calcular manuais e elétricas tornaram-se comuns. O cálculo da soma dos quadrados de uso corrente na Anova era facilmente obtido, bem como a soma de produtos que facilita, também, a análise de regressão. Entretanto, faltava qualquer capacidade de programação, só trazida pelos computadores eletrônicos que acarretaram grande economia de tempo e de mão-de-obra. Imagine fazer a inversão de uma matriz de ordem elevada ou o ajustamento de uma regressão múltipla com muitas variáveis, com uma máquina de calcular! Atualmente, um estatístico que não usa o computador é como uma espécie em extinção, cada vez mais raro de ser encontrado. Contudo, a realização de qualquer operação com um computador requer a existência de um programa apropriado, como por exemplo, o Statistical Analysis System – SAS –, o Statistical Package for Social Sciences – SPSS –, o Genstat, poderoso programa orientado primariamente para a análise de dados de experimentos planejados e para técnicas de análise multidimensional, e vários outros conhecidos pelas respectivas siglas. Os computadores são providos por uma ou mais linguagens, tais como a Fortran (Formula Translator), desenvolvida pela International Business Machines Corporation – IBM –, apropriada para trabalhos de natureza científica, e a Cobol (Common Business Oriented Language), mais usada no mundo dos negócios, por exemplo. Maiores detalhes sobre computadores podem ser vistos nos artigos de Herman Otto Hartley (1912 – 1980) em Hartley (1976), e de Nelder (1984), bem como no Capítulo 10, *Computer Programs for Survey Analysis* da 4ª edição do livro de Yates (1981). Entre outros assuntos, Hartley trata da simulação de processos estocásticos pela geração de números aleatórios, conhecida por métodos Monte

Carlo, de importante impacto na construção de modelos matemáticos. O leitor interessado nesse tema deve ler o livro de Meyer (1954).

Os “cérebros eletrônicos” – como foram chamados inicialmente os computadores – têm feito verdadeiras maravilhas, a ponto dos entusiastas da Inteligência Artificial acreditarem que, com o tempo, será possível duplicar qualquer atividade da mente humana, já que esta é também uma máquina. Entretanto, outros argumentam que o processo criativo da mente humana é de natureza diferente e jamais será reproduzido numa máquina. O uso intensivo dos computadores afastou o estatístico do escrutínio inteligente dos dados, com conseqüências maléficas, se não forem utilizados com sabedoria, pois como diz Yates “os computadores são bons serventes, mas maus mestres”. Um exemplo dado por Hartley ilustra a inspeção dos erros residuais, isto é, das divergências entre dados observados e os valores ajustados pela regressão, que é altamente vantajosa para o pesquisador aprender algo sobre seus dados. Por exemplo, o gráfico dos resíduos contra os valores calculados pela regressão linear pode indicar a falta de um termo quadrático ou de ordem superior.

Quando usados judiciosamente, os pacotes estatísticos têm ajudado de modo extraordinário tanto os estatísticos como os pesquisadores. A lição aprendida foi não se intimidar em tratar grandes massas de dados, o que levou a uma nova onda de atividades conhecida por análise exploratória de dados, cujo expoente máximo foi John Wilder Tukey (1915 – 2000). Suas contribuições sobre o assunto encontram-se principalmente no artigo de Tukey (1962) e no seu livro, Tukey (1977). Para Tukey, os problemas da ciência e suas aplicações tecnológicas, incluindo entre estas a engenharia, a agricultura e a

medicina, não iniciam nem terminam com respostas ordenadas, daí a reabilitação da estatística descritiva, começando com análises gráficas e visuais.

A análise exploratória, com ênfase nos aspectos descritivos, não elimina a análise confirmatória, de cunho inferencial, mas se completam, como escreve Tukey (1980). Para os estatísticos aplicados, Tukey é conhecido por seu teste para comparar todo e qualquer contraste entre duas médias, baseado na amplitude total “estudentizada” (studentized range), chamada na literatura, de teste de Tukey, cuja aplicação requer tabela especial, encontrada no livro de Snedecor (1937), citado na Seção 3, a partir da quinta edição.

O processo de procurar valiosas informações em enormes massas de dados é conhecido por mineração de dados (data mining), cujo exemplo é o projeto do genoma humano, que já armazenou centenas de gigabytes de dados. A mineração de dados é considerada um assunto interdisciplinar, que representa a confluência de várias idéias, inclusive da análise exploratória de dados, entre outras. Seu objetivo principal é encontrar estrutura nos dados, distinguindo-se da Estatística pela maior ênfase em algoritmos. Aliás, foi tirando proveito das redes neurais e dos algoritmos genéticos que foram desenvolvidos os métodos de análise de dados baseados no aprendizado de máquinas (machine learning). Sobre esse assunto é aconselhável a consulta ao livro de Hand et al., (2000). Esse e outros assuntos relacionados são também tratados no artigo de Rao (1999), baseado em sua palestra proferida na abertura da Sexta Escola de Modelos de Regressão, em 8 de fevereiro de 1999, em Brasília, de suma importância.

Segundo Rao, as limitações dos atuais métodos estatísticos em tratar grandes massas de dados, levaram os cientistas da computação, engenheiros e aqueles que trabalham em pesquisa operacional, a sugerir a utilização da mineração de dados para esse fim.

A pesquisa operacional teve início na Grã-Bretanha, durante a Segunda Guerra Mundial, num esforço científico interdisciplinar para resolver problemas militares, mas depois recebeu considerável desenvolvimento nos Estados Unidos da América. Hoje, a pesquisa operacional é descrita como a aplicação do método científico a problemas de decisão e gerenciamento nas áreas de negócios, indústria e administração. Seus elementos essenciais incluem o desenvolvimento de um modelo quantitativo da operação estudada e da verificação e refinamento do modelo por meio de observações quantitativas e experimentos controlados. Assim, ela proporciona aos administradores e executivos uma base quantitativa para a tomada de decisões, lidando muitas vezes com problemas de interesse estatístico. Contudo, enquanto na análise estatística raciocinam-se fatos observados no mecanismo que os geraram, na pesquisa operacional são usados modelos matemáticos que deduzem fenômenos a serem comparados com os fatos observados. A formulação dos inventários como os problemas de filas (queues) e a avaliação de projetos e revisão de técnicas, conhecida pela sigla Pert (project evaluation and review technique), são exemplos de atividades relacionadas com a pesquisa operacional.

Em seu artigo, Rao menciona também que na segunda metade do século passado, houve uma mudança na pesquisa estatística, baseada preponderantemente em modelos, para a utilização de métodos não-paramétricos,

aplicáveis em amostras oriundas de qualquer distribuição, e para métodos paramétricos robustos, não influenciados por valores atípicos (outliers) ou pela contaminação de dados. Embora de rápida implementação, os métodos não-paramétricos, baseados em estatísticas de posto (rank), não tinham a eficiência dos métodos paramétricos. Receberam especial atenção os estimadores robustos da classe M, assim chamados por serem uma generalização da máxima verossimilhança, introduzidos por Peter J. Huber. A esse respeito devem ser lidos o artigo de Huber (1964) e seu livro (1981). Os recentes desenvolvimentos dos métodos bootstrap e jackknife têm tido mais êxito do que os métodos de posto, pois não se baseiam em modelos e utilizam mais informações. Contudo, sua justificação é também baseada em resultados assintóticos, para usar as próprias palavras de Rao.

Os métodos bootstrap e jackknife (esses termos são intraduzíveis), embora tenham uma feição de mineração de dados, aproximam-se mais do raciocínio estatístico, nada mais sendo que técnicas de reamostragem, atualmente de uso nos pacotes mais comuns. Para Bradley Efron (1938 –), professor de Estatística na Universidade de Stanford, Estados Unidos – provavelmente quem mais escreveu sobre esse assunto –, o método bootstrap é simples e direto, para calcular valores aproximados tais como, de vieses, erros padrões, intervalos de confiança, em quase todo problema de estimação não-paramétrica e entre os métodos genuinamente não-paramétricos é o de melhor execução. Com R. Tibishirani, ele é autor de um livro (Efron & Tibishirani, 1993) e de vários artigos sobre o assunto, dos quais serão citados: Efron (1979, 1981); Efron & Gong (1983); Efron & Tibishirani (1986) e DiCiccio & Efron (1996). O estudo do método bootstrap

e de outros métodos de reamostragem é considerado um dos desenvolvimentos mais significativos da metodologia estatística da década de 80.

Uma revisão bibliográfica extensa sobre a técnica de jackknife desde a sua introdução por M. H. Quenouille e J. W. Tukey encontra-se no artigo de Miller (1974), o qual trata, também, da redução do viés e da estimação intervalar robusta, dois diferentes usos dessa técnica. Quenouille (1949) aplicou-a originalmente como técnica não-paramétrica, na estimação do viés e Tukey (1958), na estimativa do erro-padrão. Nesse caso, a estimativa

do jackknife é $\sqrt{\frac{n}{n-1}}$ vezes a estimativa do bootstrap.

O jackknife é considerado como uma aproximação do bootstrap pela série de Taylor; este último funciona satisfatoriamente numa grande variedade de problemas de estimação. Por exemplo, ele estima corretamente, de modo assintótico, a variância da mediana, enquanto o jackknife não funciona nesse caso. A técnica de jackknife também não serve para corrigir valores atípicos. Aos interessados na leitura desse assunto, em língua portuguesa, recomenda-se que consultem o Capítulo 8 – *Métodos Empíricos* – do livro da autoria de Souza (1998), professor titular da Universidade de Brasília – UnB –, publicado pela Embrapa.

Nos últimos tempos, os métodos bayesianos tornaram-se mais usados em diversas áreas, inclusive na econometria. A esse respeito deve ser visto o livro de Zellner (1971). Já na década de 70, Francis John Anscombe (1918 – 2001), em seu artigo (1961), achava que nos anos vindouros, haveria uma revolução na metodologia estatística e uma justa apreciação só poderia ser obtida com o estudo dos dois métodos (ortodoxo e

bayesiano), numa variedade de problemas. Contudo, a maioria das análises estatísticas continua sendo feita pelos métodos freqüencistas, cujas soluções são relativamente mais simples, segundo Efron (1986). O assunto é também tratado por Efron (1998) em *R. A. Fisher in the 21st Century*, onde ele afirma que a inferência fiducial de Fisher aproximava-se do bayesianismo objetivo, no qual o elemento subjetivo é removido da escolha da distribuição a priori. Fisher e Jeffreys, autores do livro *Theory of Probability* (1961), eram virtualmente idênticos em seu objetivo, segundo David R. Cox, embora diferissem, naturalmente, em sua matemática. Sobre esse assunto devem ser lidos os artigos de Cox (1958a; 1978).

Para completar, deve ser incluída a leitura do artigo de Cornfield (1969). Jerôme Cornfield (1912 – 1979), conhecido por seus trabalhos aplicados à epidemiologia e ensaios clínicos, é bayesiano subjetivista como De Finetti, autor do livro *Teoria delle Probabilità* (1970), (traduzido para o inglês e o alemão), famoso subjetivista cujo artigo sobre o assunto De Finetti (1974) foi criticado por Joseph Berkson em *My encounter with neo-Bayesianism*, (1977). Nesse artigo, Berkson conclui que os métodos bayesianos podem ser validamente aplicados, quando a informação a priori baseia-se em premissas de evidência objetiva, mas não a problemas científicos, quando refletem uma idéia subjetiva da probabilidade, caso em que a verificação empírica é um requisito fundamental. Para finalizar, deve ser citada a opinião de Dennis Victor Lindley (1923 –), uma das maiores autoridades no assunto, autor do livro *Introduction to Probability and Statistics from a Bayesian Viewpoint – Part 1. Probability, Part 2. Inference* (1965), e de um mais popular, Lindley (1971), que assim se expressa na página 70, da Part 2, Inference:

“O enfoque Bayesiano e o ortodoxo (frequencista clássico) se complementam e juntamente fornecem um entendimento substancialmente melhor da estatística do que isoladamente”.

Referências



ALLAN, F. E.; WISHART, J. A method of estimating the yield of a missing plot in field experimental work. **Journal of Agricultural Science**, Cambridge, v. 20, p. 417-439, 1930.

ANSCOMBE, F. J. Bayesian statistics. **The American Statistician**, Washington, DC, v. 15, n. 1, p. 21-24, 1961.

BARNETT, V. **Comparative statistical inference**. New York: Wiley, 1973.

BARTLETT, M. S. R. A. Fisher and the last fifty years of statistical methodology. **Journal of the American Statistical Association**, Washington, DC, v. 60, p. 395-409, 1965.

BAYES, T. An essay towards solving a problem in the doctrine of chances. **Philosophical Transactions of the Royal Society**, v. 53, p. 370-418, 1763. Reproduzida em **Biometrika**, London, v. 45, p. 293-315, 1958.

BEHRENS, W. Ein Beitrag zur Fehlerberechnung bei weniger Beobachtungen. **Landwirtschaftliche Jahrbucher**, Berlin, v. 68, p. 87-837, 1929.

BERKSON, J. My encounter with neo-Bayesianism. **International Statistics Review**, Voorburg, The Netherlands, v. 45, p. 1-8, 1977.

BOSE, R. C. On the application of the properties of Galois fields to the problem of hyper-graeco-latin squares. **Sankhyā**, Calcutta, v. 3, p. 323-338, 1938.

BOSE, R. C.; SHRIKVANDE, S. S. On the falsity of Euler's conjecture about the non-existence of two orthogonal latin squares of order $4t + 2$. **Proceedings of the National Academy of Sciences of USA**, Washington, DC, v. 45, p. 734-737, 1959.

BOWLEY, A. L. Working-class households in Reading. **Journal of the Royal Statistical Society**, London, 1913.

BOX, G. E. P. The exploration and exploitation of response surfaces: some considerations and examples. **Biometrics**, Washington, DC, v. 10, p. 16-60, 1954.

BOX, J. F. R. A. Fisher: **the life of a scientist**. New York: Wiley, 1978.

BOX, G. E. P.; WILSON, K. P. On the experimental attainment of optimum conditions. **Journal of the Royal Statistical Society, B**, London, v. 13, p. 1-45, 1951.

BUCKLAND, W. L. A review of the literature of systematic sampling. **Journal of the Royal Statistical Society, B**, London, v. 13, p. 268-215, 1951.

COCHRAN, W. G. The distribution of quadratic forms in a normal system, with applications in the analysis of

covariance **Cambridge Philosophical Society**,
Cambridge, v. 30, n. 2, p. 179-191, 1934.

COCHRAN, W. G. **Sampling techniques**. New York:
Wiley, 1953.

COCHRAN, W. G. Fisher and the analysis of variance.
In: S. S. FIENBERG, S. S.; HINKLEY, D. V. (Ed.). R.
A. Füller. **Fisher: an appreciation**. New York: Springer
Verlag, 1980. p. 17-34.

COCHRAN, W. G.; COX, G. M. **Experimental designs**,
New York: Wiley, 1950.

CORNFIELD, J. The Bayesian outlook and its
applications. **Biometrics**, Washington, DC, v. 25, p. 617-
657, 1969.

COX, D. R. Some problems connected with statistical
inference. **Annals of Mathematical Statistics**, Washing-
ton, DC, v. 29, p. 357-372, 1958a.

COX, D. R. The regression analysis of binary sequences.
Journal of the Royal Statistical Society, B, London,
v. 20, p. 215-242, 1958b.

COX, D. R. **Planning of experiments**. New York: Wiley,
1958c.

COX, D. R. **The analysis of binary data**1970. London:
Methuen, 1970.

COX, D. R. Regression models and life tables. **Journal
of the Royal Statistical Society, B**, London, v. 34, p.
187-220, 1972.

COX, D. R. Foundations of statistical inference: the case for eclecticism. **Australian Journal of Statistics**, Canberra, v. 20, p. 43-59, 1978.

COX, D. R. The current position of statistics: a personal view. **International Statistics Review**, Voorburg, The Netherlands, v. 65, n. 3, p. 261-290, 1997.

COX, D. R.; HINKLEY, D. V. **Theoretical statistics**. London: Chapman and Hall, 1974.

COX, D. R.; MILLER, H. D. **The theory of stochastic processes**. London: Methuen, 1965.

COX, D. R.; OAKES, D. **Analysis of survival data**. London: Chapman and Hall, 1980.

COX, D. R.; REID, N. **The theory of the design of experiments**. London: Chapman and Hall, 2000.

DALENIUS, T. Advances in sample survey theory and methods. **Annals of Mathematical Statistics**, Washington, DC, v. 33, p. 325-349, 1962.

DE FINETTI, B. **Teoria delle probabilità**. Torino: Einaudi, , 1970.

DE FINETTI, B. Bayesianism: its unifying role for both the foundations and applications of statistics. **International Statistics Review**, Voorburg, The Netherlands, v. 42, p. 117-130, 1974.

DEMING, W. E. **Some theory of samplig**. New York: Wiley, 1950.

DICICCIO, T. J.; EFRON, B. Bootstrap confidence intervals. **Statistical Science**, Hayward, CA, 11, v. 30, p. 189-228, 1996.

DROESBEKE, J. J.; TASSI, P. **Les sondages**. Paris: Presses Universitaires de France, 1990. Chap. 4. Histoire de la statistique. Col. Que sais-je?

EDWARDS, A. W. F. The history of likelihood. **International Statistics Review**, Voorburg, The Netherlands, v. 42, p. 9-15, 1974.

EFRON, B. Bootstrap methods: another look at the jackknife. **The Annals of Statistics**, Hayward, CA, v. 7, p. 1-26, 1979.

EFRON, B. Nonparametrics estimates of standard error: the jackknife, the bootstrap and other methods. **Biometrika**, London, v. 68, p. 589-599, 1981.

EFRON, B. Why isn't everyone a Bayesian? **The American Statistician**, Washington, DC, v. 40 n. 1, p. 1-11, 1986b.

EFRON, B. R. A. Fisher in the 21 st century. **Statistical Science**, Hayward, CA, v.13 n. 2, p. 95-122, 1998.

EFRON, B.; GONG, G. A leisurely look at the bootstrap, the jackknife and cross-validation. **The American Statistician**, Washington, DC, v. 37, p. 36-48, 1983.

EFRON, B.; TIBISHIRANI, R. Bootstrap methods for standard errors, confidence intervals and other methods of statistical accuracy. **Statistical Science**, Hayward, CA, v. 1 n. 1, p. 54-77, 1986a.

EFRON, B; TIBISHIRANI, R. **An introduction to the bootstrap**. New York: Chapman and Hall, 1993.

EINSENHART, C. The assumptions underlying the analysis of variance. **Biometrics**, Washington, DC, v. 3, p. 1-21, 1947.

EINSENHART, C. On the transition from Student's z to Student's t. **The American Statistician**, Washington, DC, v. 33 n. 1, p. 6-10, 1979.

ELDERTON, W. P. **Frequency curves and correlation**. 2nd ed., London: Charles & Edwin Layton, 1927.

FINNEY, D. J. The fractional replication of factorial arrangements. **Annals of Eugenics**, London, v. 12, p. 291-301, 1945.

FISHER, R. A. On the absolute criterion for fitting frequency curves. **Messenger of Mathematics**, London, v. 41, p. 115-160, 1912.

FISHER, R. A. Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. **Biometrika**, London, v. 10, p. 507-521, 1915.

FISHER, R. A. The correlation between relatives on the supposition of Mendelian inheritance. **Transactions of the Royal Society of Edinburgh**, Edinburgh, v. 52, p. 399-433, 1918.

FISHER, R. A. On the probable error of a coefficient of correlation deduced from a small sample. **Metron**, Roma, v. 1, p. 1-32, 1921.

FISHER, R. A. On the interpretation of chi-square from contingency tables, and the calculation of P. **Journal of the Royal Statistical Society**, London, v. 85, p. 87-94, 1922a.

FISHER, R. A. On the mathematical foundations of theoretical statistics. **Philosophical Transactions of the Royal Society of London, A**, London, v. 222, p. 309-368, 1922b.

FISHER, R. A. The distribution of the partial correlation coefficient. **Metron**, Roma, v. 3, p. 329-332, 1924.

FISHER, R. A. Application of Student's distribution. **Metron**, Roma, v. 5, p. 90-104, 1925a.

FISHER, R. A. **Statistical methods for research workers**. Edinburgh: Oliver & Boyd, 1925b.

FISHER, R. A. The theory of statistical estimation. **Proceedings of the Cambridge Philosophical Society**, Cambridge, v. 22, p. 701-725, 1925c.

FISHER, R. A. The arrangement of field experiments. **Journal of the Ministry of Agriculture**, London, v. 33, p. 503-513, 1926.

FISHER, R. A. Inverse probability. **Proceedings of the Cambridge Philosophical Society**, Cambridge, v. 26, p. 528-535, 1930a.

FISHER, R. A. **The genetical theory of natural selection**. Oxford: Clarendon Press, 1930b.

FISHER, R. A. The logic of inductive inference. **Journal of the Royal Statistical Society**, London, v. 98, p. 39-82, 1935a.

FISHER, R. A. The fiducial argument in statistical inference. **Annals of Eugenics**, London, v. 6, p. 391-398, 1935b.

FISHER, R. A. **The design of experiments**. Edinburgh: Oliver and Boyd, 1935c.

FISHER, R. A. Statistical methods and scientific induction. **Journal of the Royal Statistical Society, B**, London, v.17, p. 69-78, 1955.

FISHER, R. A. **Statistical methods and scientific inference**. Edinburgh: Oliver and Boyd, 1956.

FISHER, R. A.; MACKENZIE, W. A. Studies in crop variation II. The manurial response of different potato varieties. **Journal of Agricultural Science**, Cambridge, v. 13, p. 311-320, 1923.

FISHER, R. A.; WISHART, J. **The arrangement of field experiments and the statistical reduction of the results**. Harpenden: Imperial Bureau of Soil Science., 1930. (Technical Communication, 10).

FISHER, R. A.; YATES, F. **Statistical tables for biological, agricultural and medical research**. Edinburgh: Oliver and Boyd, 1938.

FRANKEL, M.; KING, B. A conversation with Leslie Kish. **Statistical Science**, Hayward, CA, v. 11, n. 1, p. 65-87, 1996.

GINI, C. Une application de la méthode représentative aux matériaux du dernier recensement italien (1^{er} décembre 1921) **Bulletin of the International Statistical Institute**, Voorburg, The Netherlands, v. 23, Liv. 2, p. 198-215, 1928.

GINI, C.; GALVANI, L. Di una applicazione del metodo rappresentativo all'ultimo censimento italiano della popolazione (1^o dicembre 1921). **Annali di Statistica**, Roma, Series 6, v. 4, p. 1-107, 1929.

GOODMAN, R.; KISH, L. Controlled selection: a technique in probability sampling. **Journal of the American Statistical Association**, Washington, DC, v. 45, p. 350-372, 1950.

GOWER, J. C. Fisher's optimal scores and multiple correspondence analysis. **Biometrics**, Washington, DC, v. 46, p. 947-968, 1990.

HAND, D. J.; MANILA, H.; SMITH, P. **Principles of data mining**. Cambridge, MA: MIT Press, 2000.

HANSEN, M. H. Some history and reminiscences on survey sampling. **Statistical Science**, Hayward, CA, v. 2, n. 2, p. 180-190, 1987.

HANSEN, M. H.; HURVITZ, W. N. On the theory of sampling from finite populations. **Annals of Mathematical Statistics**, Washington, DC, v. 14, p. 332-362, 1943.

HANSEN, M. H.; HURVITZ, W. N.; MADOW, W. G. **Sample survey methods and theory**. New York: Wiley, 1953. 2 v.

HANSEN, M. H.; MADOW, W. G. Some important events in the historical development of sampling surveys. In: OWEN, D. B. (Ed.). **On the history of statistics and probability**. New York: Marcel Dekker, 1976. p. 75-102.

HANSEN, M. H.; STEINBERG, J. Control of errors in surveys. **Biometrics**, Washington, DC, v. 12, p. 462-474, 1956.

HARTLEY, H. O. The impact of computers on statistics. In: OWEN, D. B. (Ed.). **On the history of statistics and probability**. New York: Marcel Dekker, 1976. p. 421-442.

HORVITZ, D. G.; THOMPSON, D. J. A generalization of sampling without replacement from a finite population. **Journal of the American Statistical Association**, Washington, DC, v. 47, p. 663-685, 1952.

HUBER, P. J. Estimation of location parameters. **Annals of Mathematical Statistics**, Washington, DC, v. 25, p. 73-101, 1964.

HUBER, P. J. **Robust estimation**. New York: Wiley, 1981.

IRWIN, J. O. Mathematical theorems involved in the analysis of variance. **Journal of the Royal Statistical Society, A**, London, v. 94, p. 284-300, 1931.

JEFFREYS, H. **Theory of probability**. 3rd ed. Oxford: Clarendon Press, 1961.

JESSEN, R. J. Statistical investigation of a sample survey for obtaining farm facts. **Iowa Agricultural Experiment Station Research Bulletin**, Ames, IA, v. 304, p. 1-104, 1942.

KEMPTHORNE, O. A simple approach to confounding and fractional replication in factorial experiments. **Biometrika**, London, v. 34, p. 225-272, 1947.

KEMPTHORNE, O. **The design and analysis of experiments**. New York: Wiley, 1952.

KING, A. J.; JESSEN, R. J. The master sample of agriculture. **Journal of the American Statistical Association**, Washington, DC, v. 40, p. 38-56, 1945.

KISH, L. **Survey sampling**. New York: Wiley, 1965.

KRUSKAL, W.; MOSTELLER, F. Representative sampling, IV. The history of the concept in statistics, 1895-1939. **International Statistics Review**, Voorburg, The Netherlands, v. 48, p. 169-185, 1980.

LEHMANN, E. L. **Testing statistical hypotheses**. New York: Wiley, 1959.

LEHMANN, E. L. **Theory of point estimation**. New York: Wiley, 1983.

LINDLEY, D. V. **Introduction to probability and statistics from a Bayesian viewpoint. Part I. Probability Part II Inference**. Cambridge University Press, 1965.

LINDLEY, D. V. **Making decisions**. New York: Wiley, 1971.

MADOW, W. G. **Teoria dos levantamentos por amostragem**. Rio de Janeiro: IBGE: Conselho Nacional de Estatística, 1951.

MADOW, W. G.; MADOW, L. H. On the history of systematic sampling. **Annals of Mathematical Statistics**, Washington, DC, v. 15, p. 1-24, 1944.

MAHALANOBIS, P. C. On large-scale sample surveys. **Philosophical Transactions of the Royal Society of London, B**, London, v. 231, p. 329-451, 1944.

MAHALANOBIS, P. C. Recent experiments in statistical sampling in the Indian Statistical Institute. **Journal of the Royal Statistical Institute, A**, Calcutta, v. 109, p. 326-370, 1946.

MEYER, H. A. **Symposium on Monte Carlo methods**. New York: Wiley, 1954.

MILLER, R. G. The jackknife. **Biometrika**, London, v. 61, p. 1-15, 1974.

MOSER, C. A. The use of sampling in Great Britain. **Journal of the American Statistical Association**, Washington, DC, v. 44, p. 231-259, 1949.

MOSER, C. A.; STUART, A. An experimental study of quota sampling. **Journal of the Royal Statistical Society, A**, London, v. 116, p. 349-405, 1953.

MOSER, C. A. Recent developments in the sampling of human populations in Great Britain. **Journal of the American Statistical Association**, Washington, DC, v. 50, p. 1195-1214, 1955.

MOSER, C. A. **Survey methods in social investigation**. London: Heineman, 1961.

NELDER, J. A. Present position and potential developments: some personal views – statistical computing. **Journal of the Royal Statistical Society, A**, London, v. 147, p. 151-160, 1984.

NEUMANN, J. von; MORGENSTERN, O. **Theory of games and economic behavior**. Princeton, NJ: Princeton University Press, 1944

NEYMAN, J. On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. **Journal of the Royal Statistical Society, A**, London, v. 97, p. 558-625, 1934.

NEYMAN, J. Statistical problems in agricultural experimentation, with the collaboration of K. Iwaskiewicz e St. Kolodzieczyk Suppl. **Journal of the Royal Statistical Society**, London, v. 2 n. 2, p. 107-180, 1935.

NEYMAN, J. Contributions to the theory of sampling human populations. **Journal of the American Statistical Association**, Washington, DC, v. 33, p. 101-116, 1938.

NEYMAN, J. Fiducial argument and the theory of confidence intervals. **Biometrika**, London, v. 32, p. 128-150, 1941.

NEYMAN, J. Fisher's collected papers. **The Scientific Monthly**, Washington, DC, v. 72, p. 406-408, 1951.

NEYMAN, J. Current notes. **Journal of the Royal Statistical Society, A**, London, v. 115, p. 602, 1952 .

NEYMAN, J. Experimentation with weather control. **Journal of the Royal Statistical Society, A**, London, v. 130, p. 285-325, 1967.

NEYMAN, J. On the application of probability theory in agricultural experiments. **Statistical Science**, Hayward, CA, v. 5 n. 4, p. 463-480, 1990.

NEYMAN, J.; LE CAM, L. **Probability models and cancer**. New York: Elsevier, 1982.

NEYMAN, J.; PEARSON, E. S. On the use and interpretation of certain test criterion for purposes of statistical inference. **Biometrika**, London, v. 20, Parte 1, p. 175-240; Part 2, p. 260-294, 1928.

NEYMAN, J.; PEARSON, E. S. On the problem of the most efficient test of statistical hypotheses. **Philosophical Transactions of the Royal Society, A**, London, v. 231, p. 289-317, 1933a.

NEYMAN, J.; PEARSON, E. S. The testing of statistical hypotheses in relation to probabilities a priori. **Proceedings of the Cambridge Philosophical Society**, Cambridge, v. 29, p. 492-510, 1933b.

NEYMAN, J.; PEARSON, E.S. Contributions to the theory of testing statistical hypotheses. I. Unbiased critical

regions of type A and type A1, **Statistical Research Memoirs**, London, v. 1, p. 1-37, 1936.

PEARSON, E. S. Student as a statistician, **Biometrika**, London, v. 30, p. 211, 1938.

PEARSON, E. S. The Neyman-Pearson story (1926-1934). In: PEARSON, E. S.; KENDALL, M. G. **Studies in the history of statistics and probability**. London: Charles Griffin, 1970. p. 455-477.

PEARSON, K. Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material. **Philosophical Transactions of the Royal Society of London, A**, London, v. 186, p. 356-360, 1895.

PEARSON, K. Regression, heredity and panmixia. **Philosophical Transactions of the Royal Society of London, A**, London, v. 187, p. 253-318, 1896.

PEARSON, K. On the criterion that a given system of deviation from the possible in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. **Philosophical Magazine**, London, 5th Series, v. 50, p. 157-175, 1900.

PEARSON, K. On the correlation of characters not quantitatively measurable. **Philosophical Transactions of the Royal Society of London, A**, London, v. 195, 1901.

PLACKETT, R. L. Current trends in statistical inference. **Journal of the Royal Statistical Society, A**, London, v. 129, p. 249-267, 1966.

QUENOUILLE, M. H. Approximate tests of correlation in time series. **Journal of the Royal Statistical Society, B**, London, v. 11, p. 68-84, 1949.

RAO, C. R. General methods of analysis for incomplete block designs. **Journal the American Statistical Association**, Washington, DC, v. 42, p. 541-561, 1947.

RAO, C. R. **Advanced statistical methods in biometric research**. New York: Wiley, 1952.

RAO, C. R. Apparent anomalies and irregularities in maximum likelihood estimation. **Sankhyā, A**, Calcutta, v. 24, p. 73-162, 1962.

RAO, C. R. **Linear statistical inference and its applications**. New York: Wiley, 1973.

RAO, C. R. Statistics: A technology for the millenium. **International Journal of Mathematical and Statistics Sciences**, Brasília, v. 8, n. 1, p. 5-25, 1999.

REID, C. **Neyman-from life**. New York: Springer-Verlag, 1982.

REID, N. A conversation with Sir David Cox. **Statistical Science**, Hayward, CA, v. 9, n. 3, p. 439-455, 1994.

SAMPFORD, M. R. Methods of cluster sampling with and without replacement for clusters of unequal sizes. **Biometrika**, London, v. 49, p. 27-40, 1962.

SAVAGE, L. J. On rereading R. A Fisher. **The Annals of Statistics**, Hayward, CA, v. 4, p. 441-500, 1976.

SCHEFFÉ, H. **The analysis of variance**. New York: Wiley, 1959.

SENG, Y. P. Historical survey of the development of sampling theory and practice. **Journal of the Royal Statistical Society, A**, London, v. 114, p. 214-231, 1951.

SNEDECOR, G. W. **Statistical methods applied to experiments in agriculture and biology**. Ames: The Iowa State College Press, 1937.

SOUZA, G. da S e. **Introdução aos modelos de regressão linear e não-linear**. Brasília: Embrapa-SEA: Embrapa-SPI, 1998. 489 p.

SOUZA, J. **Pesquisa eleitoral-críticas e técnicas**. Brasília: Centro Gráfico do Senado Federal, 1990.

STEPHAN, F. F. History of the uses of modern sampling procedures. **Journal of the American Statistical Association**, Washington, DC, v. 45, p. 12-32, 1949.

STEVENS, W. L. The completely orthogonalized latin squares. **Annals of Eugenics**, London, v. 9, p. 82-93, 1938.

STIGLER, S. M. Stigler's law of eponomy. **Transactions of the New York Academy of Sciences**, New York, 2nd Series, v. 39, p. 147-157, 1980.

STIGLER, S. M. **The history of statistics: the measurement of uncertainty before 1900**. Cambridge, MA: Harvard University Press, 1986.

STUDENT. The probable error of a mean. **Biometrika**, London, v. 6, p. 1-25, 1908a.

STUDENT. Probable error of a correlation coefficient. **Biometrika**, London, v. 6, p. 302-310, 1908b.

SUKHATME, P. V. Bias in the use of small-size plots in sample surveys for yield. **Nature**, London, v. 157, n. 3393, p. 630, 1946.

SUKHATME, P. V. The problem of plot-size in large-scale field surveys. **Journal of the American Statistical Association**, Washington, DC, v. 42, p. 297-310, 1947.

SUKHATME, P. V.; SUKHATME, B. V. **Sampling theory of surveys with applications**. 2nd ed. Rome: Food and Agriculture Organization of the United Nations, 1970.

TCHUPROV, A. A. On the mathematical expectation of the moments of frequency distributions in the case of correlated observations. **Metron**, Roma, v. 2, p. 646-683, 1923.

TUKEY, J. W. Bias and confidence in not-quite large samples (abstrat) **Annals of Mathematical Statistics**, Baltimore, MD, v. 29, p. 614, 1958.

TUKEY, J. W. The future of data analysis. **Annals of Mathematical Statistics**, Baltimore, MD, v. 33, p. 1-67, 1962.

TUKEY, J. W. **Exploratory data analysis**. Reading: Addison-Wesley, 1977.

TUKEY, J. W. We need both exploratory and confirmatory. **The American Statistician**, Washington, DC, v. 34, n. 1, p. 23-25, 1980.

WALD, A. Contributions to the theory of statistical estimation and testing hypotheses. **Annals of Mathematical Statistics**, Washington, DC, v. 10, p. 299-326, 1939.

WALD, A. **Sequential analysis**. New York: Wiley, 1947.

WALD, A. **Statistical decision functions**. New York: Wiley, 1950.

YATES, F. The principles of orthogonality and confounding in replicated experiments. **Journal of Agricultural Science**, Cambridge, v. 23, p. 108-145, 1933.

YATES, F. **The design and analysis of factorial experiments**. Harpenden: Imperial Bureau of Soil Science, 1937. (Technical Communication, n. 35).

YATES, F. A review of recent statistical developments in sampling and sampling surveys. **Journal of the Royal Statistical Society, A**, London, v. 139, p. 12-43, 1946.

YATES, F. **Sampling methods for censuses, and surveys**. London: Charles Griffin, 1949.

YATES, F. The early history of experimental designs. In: SRIVASTAVA, J. N. (Ed.). **A survey of statistical designs on linear models**. Amsterdam: North Holland, 1975. p. 581-592.

YATES, F. Fiducial probability, recognizable sub-sets and Behrent's test. **Biometrics**, Washington, DC, v. 20, p. 343-360, 1964a.

YATES, F. Sir Ronald Fisher and the design of experiments. **Biometrics**, Washington, DC, v. 20, p. 312, 1964b.

YATES, F. Computers-the second revolution in statistics. **Biometrics**, Washington, DC, v. 22, p. 223-251, 1966.

YATES, F. Computer programs for survey analysis. In: **SAMPLING methods for censuses and surveys**. 4th ed. London: Charles Griffin, 1981. Cap. 10.

YATES, F.; GRUNDY, P. M. Selection without replacement from within strata with probability proportional to size. **Journal of the Royal Statistical Society, B**, London, v. 15, p. 253-261, 1953.

YATES, F.; MATHER, K. Ronald Aylmer Fisher. **Biographical Memoirs of Fellows of the Royal Society of London**, London, v. 9, p. 91-120, 1963.

YULE, G. U.; KENDALL, M. G. **An introduction to the theory of statistics**. London: Charles Griffin, 1937.

ZELLNER, A. **An introduction to Bayesian inference to econometrics**. New York: Wiley, 1971.

Títulos Lançados

Nº 1 – A pesquisa e o problema de pesquisa: quem os determina?
Ivan Sergio Freire de Sousa

Nº 2 – Projeção da demanda regional de grãos no Brasil: 1996 a 2005
Yoshihiko Sugai, Antonio Raphael Teixeira Filho, Rita de Cássia Milagres Teixeira Vieira e Antonio Jorge de Oliveira

Nº 3 – Impacto das cultivares de soja da Embrapa e rentabilidade dos investimentos em melhoramento
Fábio Afonso de Almeida, Clóvis Terra Wetzel e Antonio Flávio Dias Ávila

Nº 4 – Análise e gestão de sistemas de inovação em organizações públicas de P&D no agronegócio
Maria Lúcia D'Apice Paez

Nº 5 – Política nacional de C&T e o programa de biotecnologia do MCT
Ronaldo Mota Sardenberg

Nº 6 – Populações indígenas e resgate de tradições agrícolas
José Pereira da Silva

Nº 7 – Seleção de áreas adaptativas ao desenvolvimento agrícola, usando-se algoritmos genéticos
Jaime Hidehiko Tsuruta, Takashi Hoshi e Yoshihiko Sugai

Nº 8 – O papel da soja com referência à oferta de alimento e demanda global
Hideki Ozeki, Yoshihiko Sugai e Antonio Raphael Teixeira Filho

Nº 9 – Agricultura familiar: prioridade da Embrapa
Eliseu Alves

Nº 10 – Classificação e padronização de produtos, com ênfase na agropecuária: uma análise histórico-conceitual
Ivan Sergio Freire de Sousa

Nº 11 – A Embrapa e a aqüicultura: demandas e prioridades de pesquisa
Júlio Ferraz de Queiroz, José Nestor de Paula Lourenço e Paulo Choji Kitamura (eds.)

Nº 12 – Adição de derivados da mandioca à farinha de trigo: algumas reflexões

Carlos Estevão Leite Cardoso e Augusto Hauber Gameiro

Nº 13 – Avaliação de impacto social de pesquisa agropecuária: a busca de uma metodologia baseada em indicadores

Levon Yegianantz e Manoel Moacir Costa Macêdo

Nº 14 – Qualidade e certificação de produtos agropecuários

Maria Conceição Peres Young Pessoa, Aderaldo de Souza Silva e Cilas Pacheco Camargo

Nº 15 – Considerações estatísticas sobre a lei dos julgamentos categóricos

Geraldo da Silva e Souza

Nº 16 – Comércio internacional, Brasil e agronegócio

Luíz Jésus d'Ávila Magalhães

Nº 17 – Funções de produção – uma abordagem estatística com o uso de modelos de encapsulamento de dados

Geraldo da Silva e Souza

Nº 18 – Benefícios e estratégias de utilização sustentável da Amazônia

Afonso Celso Candeira Valois

Nº 19 – Possibilidades de uso de genótipos modificados e seus benefícios

Afonso Celso Candeira Valois

Nº 20 – Impacto de exportação do café na economia do Brasil – análise da matriz de insumo-produto

Yoshihiko Sugai, Antônio R. Teixeira Filho e Elisio Contini

Impressão e acabamento
Embrapa Informação Tecnológica

República Federativa do Brasil

Luiz Inácio Lula da Silva
Presidente

Ministério da Agricultura, Pecuária e Abastecimento

Roberto Rodrigues
Ministro

Empresa Brasileira de Pesquisa Agropecuária

Conselho de Administração

José Amauri Dimázio
Presidente

Clayton Campanhola
Vice-Presidente

Alexandre Kalil Pires
Hélio Tollini

Ernesto Paterniani

Luís Fernando Rigato Vasconcellos
Membros

Diretoria-Executiva

Clayton Campanhola
Diretor-Presidente

Gustavo Kauark Chianca
Herbert Cavalcante de Lima
Mariza Marilena T. Luz Barbosa
Diretores-Executivos

Secretaria de Gestão e Estratégia

Maria Luíza Falcão Silva
Chefe

Embrapa Informação Tecnológica

Fernando do Amaral Pereira
Gerente-Geral