VIDYA, v. 34, n. 1, p. 99-116, jan./jun., 2014 - Santa Maria, 2013. **ISSN** 2176-4603

ALGUMAS CONSIDERAÇÕES SOBRE A DENOMINADA CURVA NORMAL

SOME CONSIDERATIONS ABOUT THE NAMED NORMAL CURVE

LORÍ VIALI*

RESUMO

A curva normal ou distribuição normal tem um papel preponderante tanto na teoria da probabilidade quanto na estatística. Na probabilidade, ela representa um dos mais conhecidos modelos teóricos, cuja principal aplicação é na inferência estatística. A partir do século XIX, o modelo foi popularizado e passou e ser utilizado pelas áreas sociais e humanas. Hoje o termo normal ou sua denominação mais comum *bell curve* é utilizado tanto na literatura científica quanto na popular. Ela é conhecida e empregada por educadores, sociólogos, médicos e jornalistas entre outros. Poucos modelos matemáticos ganharam tal destaque e relevância. Esse artigo traça um panorama de seu surgimento, desenvolvimento, divulgação, caracterização e aplicações.

Palavras-chave: História da curva normal. Distribuição normal. Modelo normal.

ABSTRACT

The normal curve or normal distribution has a key role both in probability theory and in statistics. In probability it represents a major theoretical model, whose main application is in statistical inference. From the nineteenth century the model was popularized and used by humanities and social areas. Today normal term or its most common denomination bell curve is used in both the scientific literature and in popular. It is known and used by educators, sociologists, doctors and journalists amongst others. Few mathematical models have gained such prominence and relevance. This article provides an overview of its emergence, development, dissemination, characterization and applications of the model.

Keywords: *history of normal curve. normal distribution. normal model.*

* Professor Doutor. Titular do Departamento de Estatística, Faculdade de Matemática da PUCRS. Professor permanente do PPG EDUCEM da Faculdade de Física da PUCRS e professor Associado do Departamento de Estatística, Instituto de Matemática da UFRGS.

INTRODUÇÃO

A partir do início do século passado, a Estatística teve o seu grande desenvolvimento graças às contribuições da Probabilidade. Sabe-se que a Estatística Inferencial depende totalmente da Probabilidade. Mas essa harmonia e simbiose entre as duas áreas não significa que ambas são idênticas. A Probabilidade continua a existir e a se desenvolver a despeito de suas aplicações a Estatística. O mesmo ocorre com a Matemática que independente de suas inúmeras aplicações continua a crescer e a progredir por si só.

A Estatística pode ser caracterizada como a análise e a interpretação de dados, tratando-se de uma disciplina aplicada. A Probabilidade, por sua vez, é um ramo da Matemática que trata da incerteza, isto é, procura modelar fenômenos não determinísticos. A origem de ambas é diferente, assim como, cada uma trilhou o seu próprio caminho até o final do século 19 ou, talvez um pouco antes, quando ocorreu o encontro e partir daí desenvolveu-se a Inferência e as coisas não foram mais as mesmas. A Probabilidade continuou a evoluir independentemente, mas o vigor dessa união foi mais prolífico para a Estatística.

Hoje praticamente todos os campos do conhecimento possuem aplicações específicas da Estatística, sendo que alguns ganharam denominação própria como a Econometria, a Bioestatística ou Biometria, a Psicometria, a Sociometria, a Quimiometria, a Geoestatística, a Astroestatística, a Demografia e a Epidemeologia. Outros ainda emprestam o nome como a Estatística Ambiental, Industrial, Financeira, Climatológica, Computacional, Matemática, Espacial, Oficial e a Agrícola. Existem ainda áreas que aparentemente nada tem a ver, mas que em geral são incluídas dentro desse grande termo guarda-chuva "Estatística". A Simulação ou Método Monte Carlo, a Teoria da Decisão, dos Jogos e da Informação, os Ensaios Clínicos, os Números Índices, o Projeto de Experimentos, a Atuária, a Amostragem e o Controle Estatístico do Processo. É praticamente impossível dizer se é a Estatística ou a Probabilidade quem mais contribui a cada uma dessas áreas.

Tanto a Estatística quanto a Probabilidade tem-se mostrado férteis para a criação de novas áreas de conhecimento. O entendimento de cada uma dessas áreas depende do conhecimento de suas raízes. Como foi mencionado, o que ocorreu da junção da Estatística com a Probabilidade foi uma fonte de novas aplicações sem, no entanto, eliminar cada uma das duas áreas individualmente. Assim, a Probabilidade continua como um dinâmico ramo da Matemática, bastante jovem, quando comparada com a Aritmética, a Álgebra ou ao Cálculo. A Probabilidade só foi axiomatizada, em 1933, com a obra de Kolmogorov[1] "Fundamentos da Teoria da Probabilidade". Talvez seja o desenvolvimento recente o responsável por parte da confusão encontrada.

A Probabilidade pode e deve ser estudada e compreendida isoladamente, isto é, sem a Estatística. Os Parâmetros Curriculares Nacionais (PCN+) da Educação Básica percebem e destacam essa diferença quando colocam (BRASIL, 2002, p. 126): "a Estatística e a Probabilidade devem ser vistas, então, como um conjunto de ideias e procedimentos que permitem aplicar a Matemática em questões do mundo real, mais especialmente aquelas provenientes de outras áreas." No Ensino Superior, a realidade é diferente. A quase totalidade dos livros textos de Estatística apresenta um capítulo ou mais de probabilidade. Nesses capítulos, geralmente são tratados conceitos básicos, modelos probabilísticos discretos e contínuos. Dentre os modelos contínuos, a curva normal é, quase sempre, o

---

1 Andrey Nikolaevich Kolmogorov (1903 - 1987). Matemático Russo com contribuições para a estatística, matemática, probabilidade, teoria da informação, da complexidade, da turbulência, dos processos estocásticos, dos sistemas dinâmicos, dos autômatos, dos algoritmos, das equações diferenciais, entre outras.

único apresentado. Os livros mais detalhados vão um pouco além apresentando outros modelos. O modelo normal é indispensável para o desenvolvimento de assuntos da estatística como a teoria da amostragem, da estimação por intervalo e dos testes de hipóteses, para citar alguns.

RELEVÂNCIA

A curva normal, curva de Gauss, de Gauss-Moivre-Laplace ou ainda curva dos erros tem importância tanto prática quanto teórica. Sua importância prática reside no fato de que muitos fenômenos ou variáveis naturais têm um comportamento que segue esse modelo. A esse respeito DAW (1996, p. 25) coloca "parece que todos acreditam na Lei dos Erros (isto é, na distribuição normal), os pesquisadores porque acham que ela é um teorema matemático e os matemáticos porque pensam que ela é um fato experimental." Variáveis biométricas, em geral, podem ser modeladas bastante bem se utilizando essa função ou distribuição. Esse modelo desempenha um papel preponderante na inferência estatística uma vez que médias de amostras de populações normais são também normais e à medida que n (tamanho da amostra) aumenta a média de uma amostra retirada de qualquer modelo (população) tende a normalidade.

A importância teórica reside no fato de que vários outros modelos probabilísticos, tanto discretos quanto contínuos, podem ser representados pelo modelo normal quando satisfeitas algumas condições de convergência. Assim, modelos como a Binomial, a Poisson e a Hipergeométrica tendem para uma normal em condições apropriadas. Distribuições contínuas como a t (de Student), a χ^2 (de Helmert[2]) e a F (de Snedecor[3]) também convergem para a normal à medida que os respectivos graus de liberdade (gl/parâmetros) aumentam.

Pierre Simon, Marques de Laplace (*)

Pierre Simon, Marques de Laplace, nascido em 23 de março de 1749 em Beaumont-en-Auge, França, e falecido em 5 de março de 1827 em Paris, França, foi o primeiro a estudar o problema da agregação de várias observações, em 1774. Embora ele tenha dada uma solução própria derivando a que é conhecida como distribuição Laplaciana, ele foi o primeiro a obter o valor da integral $\int e^{-z^2} dz = \sqrt{\pi}$ em 1782, determinando, dessa forma, a constante de normalização[4] para a distribuição normal. Foi Laplace, também, que, em 1810, provou e apresentou à academia Francesa o Teorema Central do Limite (TCL), que enfatizava a importância teórica do modelo.

Johann Carl Friedrich Gauss (*)

A distribuição normal ficou mais conhecida como curva de Gauss[5], dando razão a lei da hiponímia de Stigler (2002, p. 277), que coloca: "nenhuma descoberta científica é batizada com o nome do seu descobridor original." Gauss, de fato, não derivou e nem estudou essa distribuição, de acordo com Stigler, mas a associou com o método dos mínimos quadrados na sua primeira publicação sobre o assunto em 1809. O próprio Gauss cita o nome de Laplace como conectado a distribuição no ano de 1774.

O matemático americano Robert Adrian (1775 – 1845) publicou, em 1809, dois trabalhos de-

2 Friedrich Robert Helmert (1843 – 1917). Geodesista alemão.
3 George Waddel Snedecor (1881 – 1974). Estatístico e Matemático Americano.
4 Na teoria da probabilidade, a constante de normalização é o valor que se deve multiplicar uma função positiva e com área finita para que a sua área total seja igual a um, isto é, para que ela seja uma função densidade de probabilidade.
5 Johann Carl Friedrich Gauss (1777 – 1855). Astrônomo, Físico e Matemático Alemão.

rivando a distribuição normal, simultaneamente e independentemente de Gauss. O trabalho passou praticamente despercebido pela comunidade científica internacional, até que foi redescoberto pelo meteorologista americano Cleveland Abbe (1838 – 1916), em 1871.

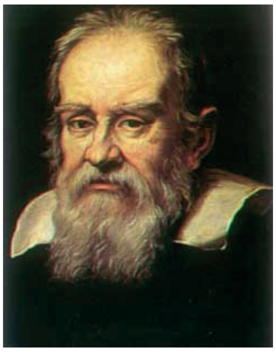
Para destacar a relevância do modelo William John Youden (1900 – 1971), engenheiro químico e estatístico americano, em sua obra *Experimentation and Measurement* coloca:

> Somente duas constantes são necessárias para determinar a altura e a largura da curva. Essas duas quantidades, que tornam possível construir a curva representando um histograma, podem ser calculadas de qualquer coleção de dados. Nós raramente temos problemas para construir a curva. Geralmente, nós fazemos uso direto das duas constantes que, em conjunto com tabelas estatísticas, são suficientes para interpretar muitos conjuntos de dados. A equação para as constantes é uma das grandes descobertas da ciência. Sua importância na interpretação de coleções de observações não deve ser subestimada.
> Esse autor tem uma pequena nota impressa como um *hobby*. Ele colocou por escrito a sua opinião sobre a importância da lei do erro normal. (YOUDEN, 1997, p. 54).

THE
NORMAL
LAW OF ERROR
STANDS OUT IN THE
EXPERIENCE OF MANKIND
AS ONE OF THE BROADEST
GENERALIZATIONS OF NATURAL
PHILOSOPHY ◆ IT SERVES AS THE
GUIDING INSTRUMENT IN RESEARCHES
IN THE PHYSICAL AND SOCIAL SCIENCES AND
IN MEDICINE, AGRICULTURE AND ENGINEERING ◆
IT IS AN INDISPENDABLE TOOL FOR THE ANALYSIS AND THE
INTEPRETATION OF THE BASIC DATA OBTAINED BY OBSERVATION AND EXPERIMENT

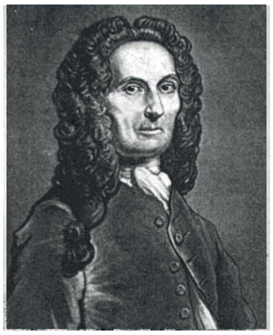
[YOUDEN, 1997, p. 54]

SURGIMENTO


Galileo Galilei (*)

A distribuição normal foi introduzida inicialmente pelo matemático francês Abraham de Moivre (1667 – 1754) em um artigo que foi reimpresso na segunda edição do seu livro “A doutrina do acaso” de 1738. Ele percebeu que à medida que o número de eventos do lançamento de moedas aumentava, a distribuição binomial se aproximava de uma curva suave. Seus resultados foram estendidos por Laplace em seu livro “Teoria Analítica da Probabilidade”, de 1812, num resultado que hoje é conhecido como Teorema de Moivre-Laplace. Laplace utilizou a distribuição na análise de erros de experimentos. Ele também mostrou que mesmo uma distribuição não sendo normal, a média de repetidas amostras dessa distribuição é aproximadamente normal e que quanto maior for o tamanho da amostra melhor será essa aproximação. O astrônomo, filósofo, físico e matemático italiano Galileo Galilei, nascido em 15 de fevereiro de 1564, em Pisa, e falecido em 8 de janeiro de 1642, em Arcetri, próximo de Florença, já havia notado que esses erros eram simétricos e que os valores pequenos apresentavam uma frequência de ocorrência maior do que os valores grandes.

O matemático Abraham De Moivre nasceu em 26 de maio de 1667 na cidade de Vitri, região da Champagne, na França. Por motivos religiosos, era Calvinista, teve que emigrar para a Inglaterra, em 1685, aos 21 anos de idade, depois de ter passado mais de dois anos na prisão em sua terra natal, devido à revogação parcial do Édito de Nantes. Na Inglaterra, para sobreviver, tornou-se professor particular e nunca conseguiu o que, de fato, almejava que era ser professor universitário de Matemática, pois os estrangeiros não tinham muitas oportunidades. A despeito disso, teve seu valor reconhecido sendo, em 1697 aos 30 anos, eleito membro da *Royal Society*. Sua obra mais conhecida, *Doctrine of Chances*, foi publicada em 1718, com reedições em 1738 e 1756 e a segunda obra *Annuities upon Lives*, lançada em 1725, teve várias edições posteriores. Na sua terceira maior obra a *Miscellanea Analytica*, publicada em Londres em 1730, apresenta uma fórmula para o cálculo aproximado de um fatorial, que foi atribuída ao matemático escocês James Stirling (1692 - 1770), que ele, depois, utilizou para obter a curva normal a partir da aproximação da distribuição binomial, em 1733. De Moivre obteve esse resultado aos 66 anos de idade e o publicou na segunda edição do seu livro a *Doctrine of Chances*, em 1738. Além disso, a obra lida com probabilidade, álgebra e trigonometria. De fato o que aconteceu é que Stirling melhorou a fórmula de De Moivre, reconhecida por ele, na edição do livro de 1738. Esse é mais um caso da lei da Eponímia. A expressão obtida para a aproximação de um valor fatorial foi: $n! \cong cn^{n+1/2}e^{-n}$, onde c é uma constante. De Moivre, contudo, não foi capaz de determinar o valor dessa constante, tarefa que foi realizada por Stirling que descobriu que $c = \sqrt{2\pi}$. Assim a fórmula que ficou conhecida como de Stirling é escrita como $n! \cong \sqrt{2n\pi}\, n^n e^{-n}$.



Abraham De Moivre (*)

De Moivre, mesmo tendo um grande talento e sendo amigo de Isaac Newton, sobreviveu apenas com os rendimentos de professor particular e morreu na pobreza em 27 de novembro de 1754, em Londres. Da mesma forma que Cardano[6], ganhou popularidade por prever o dia da própria morte. Em um raro panfleto, datado de 12 de novembro de 1733, escrito em Latim e denominado de *Approximatio ad summam terminorum binomii (a+b)^n in seriem expansi,* a distribuição normal foi derivada como uma aproximação da distribuição binomial. Segundo Pearson, “este suplemento contém o primeiro tratamento conhecido da probabilidade integral e essencialmente da curva normal.” Ele antecedeu a discussão de Laplace em meio século.” (PEARSON, 1924, p. 402; DAW, 1996, p. 2). Também segundo Daw somente duas cópias desse panfleto são conhecidas, sendo que uma está na biblioteca da Universidade de Londres e a outra em Berlim.

Ainda segundo Pearson, De Moivre publicou em 1730 o *Miscellanea Analytica* e muitas cópias desse trabalho continham o suplemento com paginação em separado que terminava em uma tabela de valores de logaritmos de fatoriais de 10 a 900, com diferenças de 10. Acrescenta que apenas poucas cópias tinham um segundo suplemento também com paginação em separado (1 a 7) e datadas de 12 de novembro de 1733. Este segundo suplemento pode somente ser adicionado às cópias vendidas três anos após o lançamento do livro original e isso contribuiu para a sua raridade.

O número 329 do *Philosophical Transactions* consistiu inteiramente de uma memória denomi-

6 Girolamo Cardano (1501 – 1576). Médico e matemático italiano. Autor da obra *Ars Magna*.

nada de *De Mensura Sortis, seu, de Probabilitate Eventuum in Ludis a Casu Fortuito Pendentibus*, cujo autor foi De Moivre (TODHUNTER, 2005, p. 136). Esse trabalho foi depois expandido e se transformou no *The Doctrine of Chances: or a Method of Calculating the Probabilities of Events in Play*, cuja primeira edição foi publicada em 1718, contendo 175 páginas. A segunda edição foi lançada em 1738 com 258 páginas e continha uma tradução para o inglês do panfleto com algumas adições. A terceira edição do livro surgiu em 1756, após a morte do autor, e continha 348 páginas.

Além da derivação da distribuição normal, um dos resultados matemáticos mais importantes obtidos por De Moivre foi o de relacionar os números complexos com a Trigonometria por meio da expressão: $(\cos x + i\,\mathrm{sen}\,x)^n = \cos(nx) + i\,\mathrm{sen}(nx)$ que foi sugerido por ele, em 1722, mas que nunca foi explicitamente expresso no seu trabalho. Esse resultado inspirou o matemático suíço Leonhard Euler (1707 – 1783) a formular que $e^{ix} = \cos x + i\,\mathrm{sen}\,x$ ou $e^{inx} = \cos(nx) + i\,\mathrm{sen}(nx)$.

Em 1967, o matemático americano Churchill Eisenhart (1913 – 1994) e o estatístico Allan Birnbaum (1923 – 1976) publicaram no periódico *The American Statistician* o artigo *Anniversaries in 1966-67 of Interest to Statisticians: Part II: Tercentennials of Arbuthnot and De Moivre*, um apanhado das contribuições de De Moivre para o Cálculo de Probabilidades (EISENHART; BIRNBAUM, 1967).

O DESENVOLVIMENTO DO MODELO

A partir de sua origem como uma aproximação da distribuição Binomial feita por De Moivre, passando por uma aproximação da distribuição Beta feita por Laplace, como solução de uma equação diferencial obtida por Gauss, em 1809, e no contexto do TCL obtido por Laplace em 1810.

Por volta de 1721, após ter lido o trabalho de Bernoulli *Ars Conjectandi*, De Moivre começou a fazer progressos no trabalho de aproximar os termos da expansão da binomial. Esse trabalho culminou, em 1733, com a publicação do que hoje é conhecida como aproximação normal da binomial. Segundo Stigler (2000), a primeira publicação sobre esse tópico apareceu em 1730 na sua *Miscellanea Analytica* ocasião em que ele já tinha quase obtido uma solução completa. Seu trabalho foi finalizado, em 1733, e publicado em Latim em uma nota curta separadamente impressa (DAW; PEARSON, 1972). Mais tarde essa nota foi traduzida para o Inglês e incluída de forma sucessivamente mais expandida na segunda e terceira edições da Doutrina das Chances.

O quinto livro da *Miscellanea Analytica* (sobre o binômio a + b elevado a potências altas) começa dando crédito ao trabalho, *Ars Conjectandi*, de James (Jakob ou Jacques) Bernoulli (1655 – 1705). A abordagem de De Moivre foi verificar o comportamento de $(1 + 1)^n$ para grandes valores de n. Em 1730, ele apresentou dois resultados, o primeiro com o termo máximo de $(1 + 1)^n$ como sendo, em notação atual, e o outro como a soma de todos os termos (isto é, 2^n), como sendo aproximadamente igual a $2A\dfrac{(n+1)^n}{n^n\sqrt{n-1}}$.

De acordo com Walker (1934, p. 73), o próprio De Moivre assim se expressou:

> Faz agora 12 anos ou mais que eu encontrei o seguinte: se o binômio 1 + 1 for elevado a uma potência bastante alta representada por n, a razão que o termo do meio possui com a soma de todos os termos, isto é, 2^n, pode ser expressa pela fração $2A\dfrac{(n+1)^n}{n^n\sqrt{n-1}}$, onde A representa o número do qual o logaritmo hiperbólico é:

$\frac{1}{12}-\frac{1}{360}+\frac{1}{1260}-\frac{1}{1680}+\cdots$, mas, em virtude de que a Quantidade $(n-1)^n$ ou $(1-1/n)^n$ estão bastante próximas para valores altos de n, o que não é difícil de provar, segue que em uma potência infinita tal Quantidade será absolutamente determinada e representa o número do qual o logaritmo hiperbólico é -1; ... e agora suponha que B representa o número do qual o logaritmo hiperbólico é $1-\frac{1}{12}+\frac{1}{360}-\frac{1}{1260}+\frac{1}{1680}-,$ tal expressão mudará para $\frac{2}{B\sqrt{n}}$. Quando iniciei o trabalho, eu fiquei satisfeito em determinar, em geral, o valor de B, que foi feito pela adição de alguns termos sobre a série acima. Contudo eu percebi que ela converge muito lentamente e vendo ao mesmo tempo que o que eu fiz respondia ao meu propósito razoavelmente, eu desisti de ir além, até que meu digno amigo Mr. James Stirling, que a meu pedido realizou o mesmo trabalho, encontrou que a Quantidade B representa a raiz quadrada da circunferência de um Círculo cujo raio é a Unidade. Assim se a circunferência for denominada por c, a Razão do Termo Médio para a Soma de todos os Termos será expressa por $\frac{2}{\sqrt{nc}}$.

De Moivre, em 1733, notou que a descoberta de Stirling não era necessária, mas que trazia elegância para a sua solução. Com a constante de Stirling, e aproximando $\left(1-\frac{1}{n}\right)^n$ por e^{-1} e $\sqrt{n-1}$ por $\sqrt{n}$, ele concluiu que a razão do termo máximo de $(1+1)^n$ por 2^n era: $\frac{2}{\sqrt{2n\pi}}$.

O que De Moivre encontrou, considerando uma notação atualizada, foi uma aproximação para $P(X = n/2)$, onde X tem uma distribuição binomial simétrica (n tentativas com n par) (STIGLER, 2000, p. 73).

Em 1739 ele levou esse resultado um pouco além, calculando qual a razão entre um termo afastado do termo do meio por um intervalo l e o termo do meio. O que ele verificou foi que:

$$\ln\left[\frac{P(X=\frac{n}{2}+l)}{P(X=\frac{n}{2})}\right] \cong -\frac{2l^2}{n}.$$

Essa expressão, que é a altura máxima da curva ou, nesse caso, é a ordenada na origem, uma vez que a média é zero, pode ser escrita como:

$$\frac{P(X=\frac{n}{2}+l)}{P(X=\frac{n}{2})} = \exp\left(-\frac{2l^2}{n}\right) \text{ ou } P(X=\frac{n}{2}+l) = P(X=\frac{n}{2})\exp\left(-\frac{2l^2}{n}\right).$$

O resultado acima, por sua vez, pode ser escrito como: $P(X=\frac{n}{2}+l) \cong \frac{2}{\sqrt{2n\pi}}\exp\left(-\frac{2l^2}{n}\right)$.

Substituindo o l, utilizado por De Moivre, por x e n por $4\sigma^2$, tem-se, em notação atual, o resultado encontrado como sendo:

$$y = \frac{2}{\sqrt{2n\pi}}\exp\left(-\frac{2l^2}{n}\right) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left[-\frac{1}{2}\left(\frac{x}{\sigma}\right)^2\right]$$ que é a expressão usual da curva normal quando $\mu = 0$.

Além de determinar a equação da distribuição normal, De Moivre calculou a área entre um desvio acima e um abaixo da média (nesse caso a origem dos eixos coordenados), isto é, entre $-\sigma$

e $+\sigma$ como sendo 0,682688. Esta é uma aproximação bastante boa, para um cálculo que utilizou um processo mecânico de quadraturas e considerando que o valor exato é 0,682689. Ele também determinou uma aproximação, do que hoje é denominado de erro provável, como sendo 0,707σ, que na notação por ele utilizada era $\frac{1}{4}\sqrt{2n}$ e que vale, de fato, 0,6745σ (WALKER, 1934).

A NORMAL COMO LEI DOS ERROS

Ao se medir repetidamente uma mesma quantidade, utilizando-se um instrumento, nota-se que os resultados não são os mesmos a cada repetição do experimento. As causas que interferem para que isso ocorra são muitas e geralmente incontroláveis. A incidência da luz na leitura do instrumento, a temperatura ambiente, a fadiga e personalidade do operador e outras condições atmosféricas tendem a interferir nas medidas de modo que se obtenham os diferentes resultados. Em alguns casos, os tipos de erros são conhecidos o suficiente para que correções sejam feitas, mas mesmo quando isso tenha sido feito, ainda restam variações suficientes nas observações para que elas não sejam consideradas idênticas. Isso ocorre porque existem causas de erros que são muito complexas para serem investigadas ou por que as leis de ação são desconhecidas (DAW, 1996, p. 3). Assim, em uma observação corrigida, o erro restante pode ser considerado como a soma de vários pequenos erros acidentais, alguns em uma direção e outros em outra, surgindo de diferentes causas. A distribuição normal tem sido, desde os tempos de Gauss, geralmente adotada como representando esses pequenos erros acidentais de medidas e passou a ser conhecida como a Lei dos Erros (DAW, 1996, p. 3).

Uma das mais sucintas derivações da distribuição normal a partir de hipóteses sobre erros em medições de posições de estrelas foi feita pelo matemático e astrônomo inglês John Herschel (1792 - 1871), em 1850. Para detalhes dessa demonstração, consultar Jaynes (2003, p. 702). O físico e matemático escocês James Clerk Maxwell (1831 - 1879), em 1860, apresentou uma versão tridimensional dos mesmos argumentos bidimensionais utilizados por Herschel para determinar a distribuição de probabilidade das velocidades das moléculas em um gás, que ficou conhecida entre os físicos como “a lei da distribuição Maxwelliana da velocidade.”

> Depreende-se dessa proposição que as velocidades estão distribuídas entre as partículas conforme a mesma lei que os erros estão distribuídos entre as observações na teoria do ‘método dos mínimos quadrados’. As velocidades variam de 0 a ∞, mas o número daquelas com grande velocidade é comparativamente pequeno (MAXWELL, 1860, p. 381).

As derivações de Maxwell e Herschel não fazem uso da teoria da probabilidade, mas apenas de propriedades da invariância geométrica que poderia ser aplicada da mesma forma em outros contextos (JAYNES, 2003). Já a derivação de Gauss, em 1809, faz uso explícito da intuição probabilística. Essa demonstração foi apenas uma nota em um trabalho relacionado com a astronomia. Ela poderia ter passado despercebida se Laplace não tivesse destacado o seu valor e publicado, em 1810, um extenso trabalho, despertando a atenção para a dedução de Gauss e demonstrando as muitas propriedades da normal como uma distribuição amostral. Desde essa época o modelo passou a ser denominado de “distribuição Gaussiana”.

De Moivre descobriu ou deduziu a curva normal, contudo, ele não percebeu sua importância e não apresentou nenhuma aplicação para a sua descoberta. Entretanto, ao longo do século XVIII, ela foi de grande valor como a lei dos erros. Em 1924, o estatístico inglês Karl Pearson publicou na revista *Biometrika*, criada por ele o artigo *Historical Note on the Origin of the Normal Curve of Errors* (PEARSON, 1934). O economista e filósofo anglo-britânico Francis Ysidro Edgeworth (1845 – 1926), professor de economia da Universidade de Oxford, fez uma palestra na *Royal Statistical Society* em 19 de junho de 1906. A apresentação teve o título de *The Generalised Law of Error, or Law of Great Numbers* e foi publicada no *Journal of the Royal Statistical* em setembro de 1906 (EDGEWORTH, 1906). O astrônomo, atuário e matemático dinamarquês, Thorvald Nicolai Thiele (1838 - 1910) publicou em 1931 um artigo no periódico *The Annals of Mathematical Statistics* denominado *Laws of Errors* (THIELE, 1931).

DIVULGAÇÃO E UTILIZAÇÃO


Lambert Adolphe Jacques Quetelet (*)

A primeira pessoa a aplicar a distribuição normal na área social foi o belga Lambert Adolphe Jacques Quetelet que coletou dados sobre medidas do peito de soldados escoceses e da altura de soldados franceses e verificou que elas podiam ser modeladas pela distribuição. Quetelet nasceu em 22 de fevereiro de 1796 em Ghent, Bélgica, e faleceu em 17 de fevereiro de 1874 em Bruxelas, Bélgica. Ele iniciou sua carreira como astrônomo, mas mudou para as ciências sociais. De acordo com Fendler e Muzaffar (2008), após Quetelet a história da curva normal é a história de como a representação matemática da probabilidade foi apropriada pelos cientistas sociais que iniciaram a descrição numérica de populações. Ele extraiu dados das medidas de 5738 peitos de soldados do *Edinburgh Medical and Surgical Journal*, de 1817, e percebeu que o padrão dessas medidas era idêntica a formada por medidas comuns em astronomia. Em outras palavras, Quetelet estava dizendo que os dados seguiam uma distribuição normal.


Adrien-Marie Legendre (*)

De forma independente, o matemático irlandês Robert Adrian (1775 - 1843), em 1808, e o alemão Johann Carl Friedrich Gauss, em 1809, desenvolveram a equação da distribuição e mostraram que ela modelava bastante bem os erros de observações astronômicas. Gauss utilizou a equação em 1809 para justificar o MMQ (Método dos Mínimos Quadrados), introduzido pelo matemático francês Adrien-Marie Legendre (1752 – 1833), em 1809, mas já utilizado por ele desde 1794 (STIGLER, 2002, p. 284).

O primeiro contato de Laplace com o modelo foi para realizar uma aproximação para uma distribuição a posteriori, sendo uma utilização diferente da que fez Gauss. Se fosse considerada apenas a conexão com o método dos mínimos quadrados, então a primeira publicação seria a de Robert Adrian, em uma revista com data de 1808, um ano antes do livro de Gauss ter sido publicado.

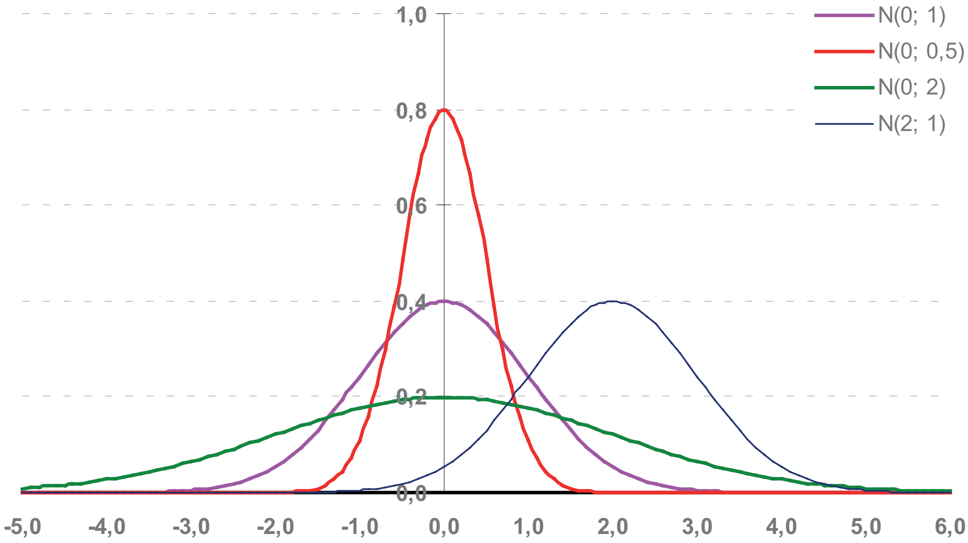
O aniversário da descoberta de De Moivre foi saudado pela educadora estatística americana Helen M. Walker (1891 – 1983), que foi professora da Universidade de Columbia de Nova Iorque, com a publicação no periódico *Journal of the American Statistical Association* do artigo *Bi-Centenary of the Normal Curve*, em 1934, comemorando os 200 anos da descoberta da distribuição (WALKER, 1934).

CARACTERÍSTICAS OU PROPRIEDADES DO MODELO

A curva normal ou modelo normal é uma família de distribuições de probabilidade que dependem dos parâmetros μ (média ou mais apropriadamente valor esperado que é o parâmetro de locação) e σ

> 0 (que nesse caso é também o desvio padrão da distribuição e é o parâmetro de escala). O modelo é simétrico em torno de seu valor esperado, isto é, $f(\mu - x) = f(\mu + x)$, e duplamente assintótico, isto é, limite quando x tende a mais infinito é igual ao limite quando x tende a menos infinito e ambos são iguais a zero. A curva possui dois pontos de inflexão, pontos onde ela muda o sinal da curvatura (a segunda derivada se anula), isto é, passa de côncava para convexa em $x = \mu - \sigma$ e em $x = \mu + \sigma$. A área útil total da curva pode ser encontrada entre $\mu - 4\sigma$ e $\mu - 4\sigma$, sendo que no intervalo $\mu - 3\sigma$ e $\mu + 3\sigma$ tem-se 99,73% da área total, entre $\mu - 2\sigma$ e $\mu + 2\sigma$ tem-se 95,44% e entre $\mu - \sigma$ e $\mu + \sigma$ tem-se 67,28%. Assim fora do intervalo compreendido entre 3 desvios padrão acima e abaixo da média pode-se encontrar apenas 0,26% da área total. Para se perceber a área compreendida abaixo ou acima de 4 desvios a contar da média, é necessário se utilizar uma aproximação superior a 4 decimais. Se X é uma variável aleatória com um comportamento normal, então é comum se utilizar a seguinte notação: $X \sim N(\mu, \sigma)$.

Figura 1 – Ilustração de algumas distribuições normais



Fonte: Construção do autor.

A respeito de a função densidade de probabilidade (fdp) ter uma expressão razoavelmente simples, o modelo normal não é integrável analiticamente, isto é, não existe uma F(x) tal que $F'(x) = f(x)$, onde $f(x) = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$. Dessa forma, é necessário o uso de tabelas ou então de recursos computacionais para a determinação das probabilidades (áreas) sob a curva. As tabelas são geralmente feitas para a variável normal padrão, que é representada por z, e determinada pela transformação $z = (x - \mu)/\sigma$, sendo simétrica em torno de zero com desvio padrão igual a 1. Assim a equação da normal padrão é dada $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$ para $-\infty < z < +\infty$. Para destacar a simetria da distribuição Normal Padrão $\phi(z)$, considerando que seu valor esperado (média) é igual a zero, pode-se escrever que $\phi(z) = \phi(-z)$. Contudo o que é, de fato, tabelada é a função acumulada de z, isto é, a função $\Phi(z) = \int_{-\infty}^{+z} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$. A função de distribuição acumulada $\Phi(z)$ é geralmente tabelada no intervalo de

-3,9 a +3,9, sendo que na maioria das tabelas no intervalo de -3,9 a -3,0 e +3,0 a +3,9 são os decis e no intervalo de -3,0 a +3,0 os percentis. É comum, também, encontrarmos tabelas apenas da parte positiva da distribuição, isto é, com x variando de 0 a 3,9. O uso da variável z permite que a fdp da normal possa ser escrita da seguinte forma: $f(x) = \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = \frac{1}{\sigma}\Phi\left(\frac{x-\mu}{\sigma}\right)$ para $x \in \Re$.

Considerando que a variável Z é simétrica em torno de zero, então se pode escrever que $\Phi(-z) = 1 - \Phi(z)$, fato esse que simplifica o cálculo de probabilidades e o uso das tabelas.

Outras propriedades incluem o coeficiente de assimetria que é obviamente zero e o coeficiente de curtose que pode ser 0 ou 3 conforme o valor for centrado ou não e a entropia que é dada por $0{,}5\ln(2\pi e\sigma^2)$. Para a curva normal padrão que apresenta variância (σ^2) igual a um, a entropia vale 1,42. A mediana e a moda são ambas iguais a μ como ocorre com as distribuições simétricas que apresentem essas três características.

NOMENCLATURA

De acordo com Stigler (2002, p. 404), a denominação "normal", para a curva $\varphi(z)$, foi surgindo gradualmente. Aparentemente, o primeiro a denominar a curva como normal foi o astrônomo, filósofo, físico e matemático americano Charles Sanders Pierce (1839 - 1914). Em um relatório publicado em 1873, ele se referiu a $\varphi(z)$ como a "curva normal dos mínimos quadrados." De forma independente, o nome foi ainda atribuído pelo antropólogo e geneticista britânico Francis Galton (1822 – 1911) e pelo economista alemão Wilhelm Lexis (1837 – 1914) por volta de 1875. A terminologia de curva em forma de sino foi cunhada em 1872 pelo francês Esprit Pascal Jouffret (1837 - 1904) que denominou a normal bivariada de superfície campanular (*bell surface*). A respeito do nome "normal", Stigler destaca que:

> Tal multiplicidade de nomes – em três países e duas línguas diferentes – é notável e certamente um sinal de grande difusão simultânea envolvendo um entendimento conceitual na década de 1870: de populações de pessoas, medidas e similares (STIGLER, 2002, p. 404).

O seu emprego generalizado é devido provavelmente a influência de Karl Pearson que declarou:

> Muitos anos atrás (em 1893) denominei a curva de Gauss-Laplace de curva normal, tal nome, enquanto evita a questão internacional da prioridade, tem a desvantagem de levar as pessoas a acreditar que todas as demais distribuições de frequência são, em um sentido ou outro, anormais (PEARSON, 1920, p. 25).

Stigler (2002, p. 405) observa, ainda que o termo "normal" não foi o único a denominar essa distribuição e que o termo "Gaussiana" é um competidor primo e atribuído em homenagem a Carl Friedrich Gauss, que, de fato, estudou a curva em 1809, mas que não se pode dizer que ele a tenha descoberto. Pode-se acrescer ainda aos dois mencionados o abominável *bell curve* (curva em forma de sino) onipresente em livros e artigos nas áreas sociais e humanas, particularmente na área educacional. O termo campanular, isto é, em forma de sino, faz sentido para a distribuição normal bidimensional, conforme atribuído por Jouffret, mas a normal unidimensional não tem um formato mais campanular do que qualquer outra de infinitas curvas unimodais, acrescenta Stigler (2002, p. 405). Ainda a esse respeito pode-se mencionar Stahl (2006, p. 96) “a estatística é a que apresenta as mais amplas aplicações de todas as disciplinas matemáticas e no centro da estatística está a distribuição normal conhecida por milhões de pessoas como a curva de sino ou a curva em forma de sino.”

Pode-se destacar que o que deve ter contribuído, principalmente nos últimos anos, para popularizar o nome de curva de sino ou curva em forma de sino foi o lançamento da obra “*Bell Curve: Intelligence and Class Structure in American Life*” de Richard J. Herrnstein e Charles Murray, em 1994, cuja capa mostrava um diagrama de uma suposta curva normal. A obra gerou uma das maiores controvérsias dos últimos anos nos meios científicos ao tocar abertamente em assuntos sensíveis como a relação entre o QI e o crime, o QI e a pobreza, nas diferenças étnicas em inteligência, taxas de fertilidade entre mulheres com diferentes níveis de inteligência e sugerir o que pode ser feito (políticas públicas) para compensar as diferenças em inteligência.

Sucedendo ao lançamento da obra de Hernstein e Murray, outros com títulos envolvendo a curva normal ou “*bell curve*”, ocorreram como “*The Bell Curve Debate*” editado por Russell Jacoby e Naomi Glauberman de 1995. “*The Bell Curve Wars: Race, Intelligence, and the Future of America*” editado por Steven Fraser de 1995. “*Poisoned Apple: The Bell-Curve Crisis and How Our Schools Create Mediocrity and Failure*” de Betty Wallace e William Graves de 1995, entre outros.

A respeito dessa popularidade sobre um nome aparentemente errôneo, uma vez que a atribuição original se referia à distribuição bidimensional (uma superfície, portanto) e assim com muito mais propriedade, poderia ser denominada de “*bell surface*”, pode-se ilustrar com uma colocação de um autor, que a exemplo de muitos outros ignoram o contexto histórico e se referem a curva normal dessa forma. Stahl (2006, p. 96) coloca “a estatística é a mais empregada de todas as disciplinas de matemática aplicada e no centro da estatística está a distribuição normal, conhecida por milhões de pessoas como a ‘*bell curve*’ ou a ‘*bell-shaped curve.*’”

APLICAÇÕES

A curva normal tem aplicações diretas servindo de modelo a muitas variáveis naturais. Alguns exemplos incluem as variáveis biométricas, tais como a altura de um grupo grande de pessoas, a taxa de colesterol, a taxa de glicemia, etc. Alguns exemplos surgem também no campo educacional e a curva normal pode modelar bastante bem conjuntos de notas de algumas disciplinas em provas de larga escala como os vestibulares e outros concursos que envolvem milhares de candidatos. Ela também é o limite de outras distribuições de probabilidade tanto discretas quanto contínuas. A distribuição binomial já foi mencionada, sendo inclusive a origem deste modelo. Variáveis contínuas que tendem a normalidade à medida que o parâmetro tende ao infinito incluem a distribuição t (de Student) e a distribuição Qui-Quadrado, entre outras. A soma de variáveis identicamente distribuídas e independentes também tende a normalidade, que é o que é enunciado pelo Teorema Central do Limite (TCL). Assim se $(X_1, X_2, ..., X_n)$ é uma amostra aleatória retirada de uma população com média μ e desvio padrão σ e se $S_n = X_1 + X_2 + ... + X_n$, então $\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le x\right) = \Phi(x)$.

Entretanto, a principal aplicação da distribuição normal, por consequência de outros modelos probabilísticos, é na inferência estatística. A inferência é a parte da estatística que se utiliza da probabilidade para generalizar resultados obtidos a partir de amostras retiradas de uma população. Frequentemente uma das hipóteses formuladas para que se possam aplicar procedimentos probabilísticos é a de que a população sob investigação tenha uma distribuição normal. Procedimentos que utilizam esse tipo de hipótese ou exigem que a população tenha um comportamento normal incluem os intervalos de confiança, os testes de hipóteses e a análise de variância (ANOVA), entre outros.

GERAÇÃO OU SIMULAÇÃO

A distribuição normal não pode ser simulada pelo método usual da inversão, pois ela não apresenta uma forma fechada (fórmula) para a distribuição acumulada (FDA). Assim, é necessário recorrer a outros procedimentos para se obter valores com comportamento normal.

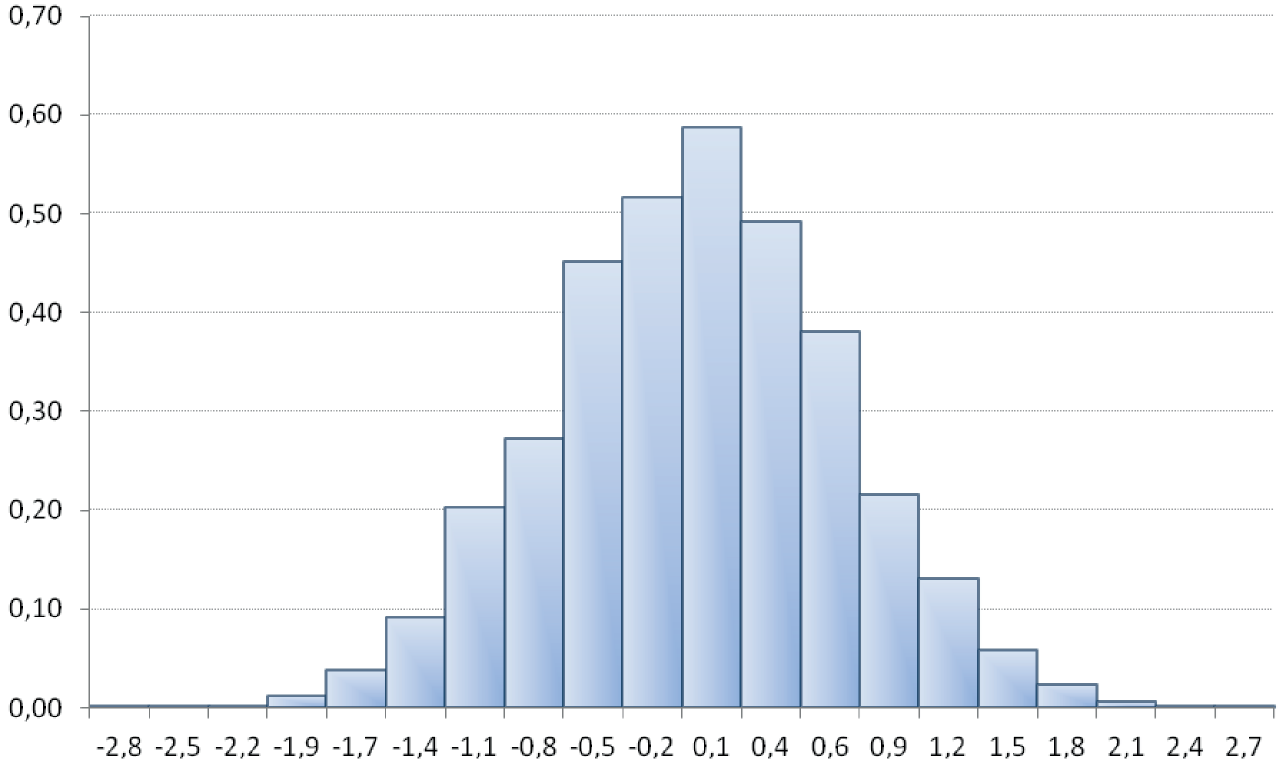
O método mais simples de simular uma distribuição normal é utilizar o TCL e gerar n valores aleatórios, isto é, variáveis uniformes U_1, U_2, ..., U_n no intervalo [0, 1] e então utilizar o Teorema Central do limite para obter $W \equiv [(U_1 + U_2 + ... + Un) - n/2] /$ que gera valores da N(0, 1). Para obter valores de uma normal qualquer, basta lembrar que $Z = (X - \mu)/\sigma$ e então $X = \sigma Z + \mu$. Nesse caso, é conveniente tomar $n = 12$, que simplifica a expressão para a geração de uma N(0, 1). Assim se $n = 12$, tem-se $W \equiv \sum_{i=1}^{12} U_i - 6$ como gerador.

Contudo, convém lembrar que esse é um método aproximado, isto é, ele converge para a N(0, 1) à medida que n tende ao infinito. Assim, outros métodos foram criados e não dependem de convergência para se obter valores da distribuição normal. Um dos métodos mais populares é denominado de Box/Müller, desenvolvido por George Edward Pelham Box (1919 – 2013) e Mervin Edgar Müller (professor emérito da Universidade estadual de Ohio), em 1958, e tem como base o seguinte resultado:

sejam U_1 e U_2 dois valores aleatórios, isto é, valores da Uniforme no intervalo [0, 1]. Então é possível mostrar que $R_1 = \cos(2\pi U_1)\sqrt{-\ln(U_2)}$ ou $R_2 = \mathrm{sen}(2\pi U_1)\sqrt{-\ln(U_2)}$ são valores de uma distribuição N(0, 1) (BOX; MULLER, 1958, p. 1).

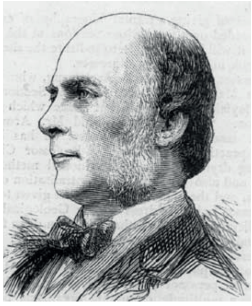
A geração da normal padrão permite a geração de outros modelos que convergem, mediante certas situações, para a normalidade como a distribuição Qui-Quadrado e a distribuição t de Student, por exemplo. A geração de valores aleatórios de modelos probabilísticos é denominada de Método de Monte Carlo, conforme nomeado por Nicholas Constantine Metropolis (1915 - 1999) e permite a solução, por simulação, de problemas que não podem ser tratados analiticamente (METROPOLIS, 1987).

Figura 2 – Histograma de 5000 valores da Normal Padrão gerados pelo método de Box/Muller.



Fonte: Construção do autor.

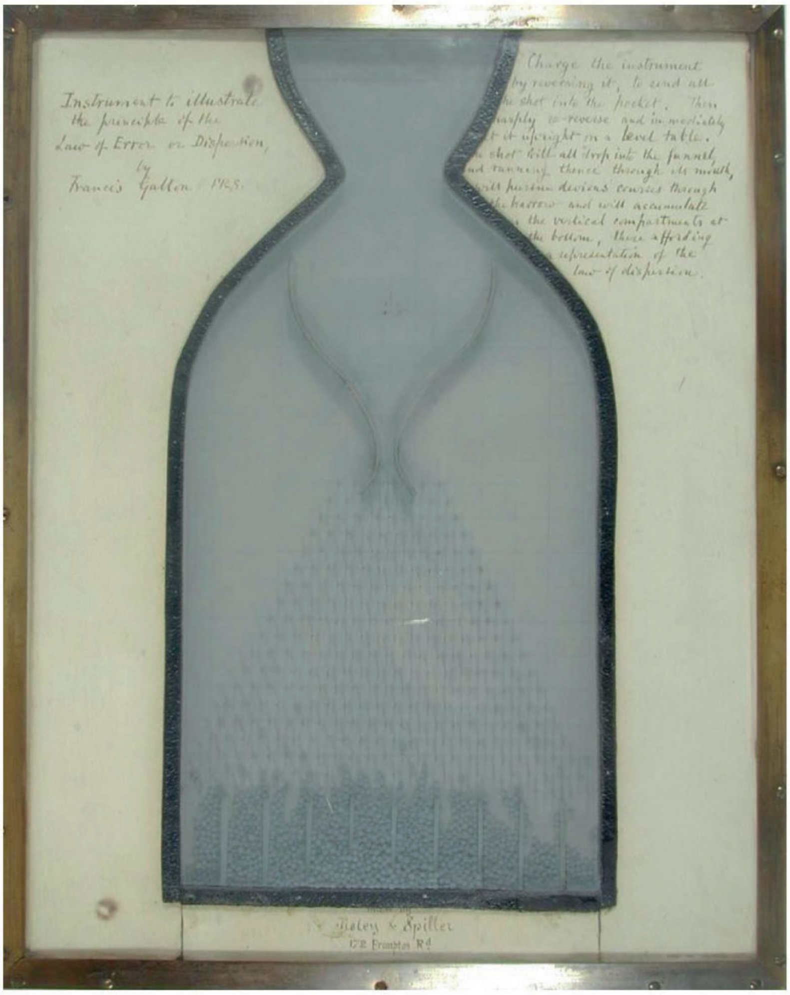
GALTON E O QUINCUNX

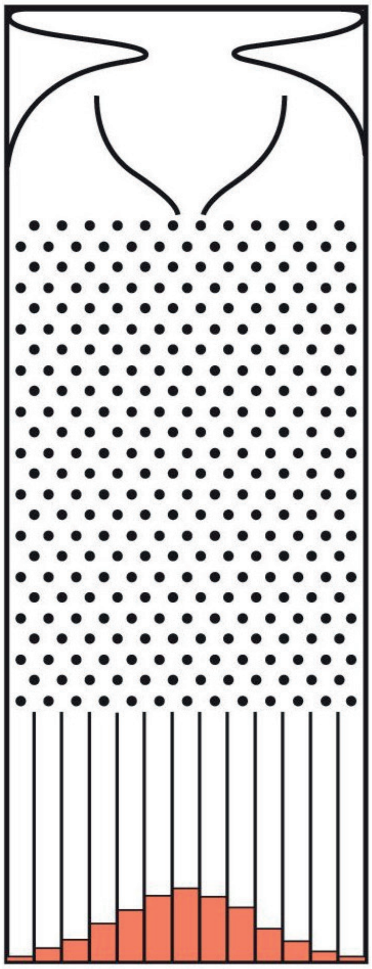


Francis Galton (*)

Além da simulação digital, a normal é, talvez, o único modelo de probabilidade que dispõe de um simulador físico. Tal simulador é um aparelho construído na Inglaterra, no século XIX. O aparelho criado, por um comerciante, para Galton, em 1873, foi nomeado de Quincunx. Ele consistia de um tabuleiro com uma espécie de funil no topo, onde eram introduzidas bolinhas que rolavam tabuleiro abaixo, batendo aleatoriamente em colunas de pinos dispostos ao longo do tabuleiro. Na parte inferior, existe uma série de compartimentos onde as bolinhas vão parar. A ideia de Galton era simular a distribuição normal. De fato, o arranjo das bolinhas segue uma distribuição Binomial, uma vez que é fruto de várias variáveis de Bernoulli que são obtidas a cada encontro da bolinha com um dos pinos. O resultado final é uma figura semelhante a um histograma que é similar a uma curva normal. Conforme Kunnert, Montag e Pöhlmann (2001), a paixão de Galton era a distribuição normal e por isso ele solicitou à companhia Tisley & Spiller que produzisse o Quincunx que ele utilizou como um recurso de ensino, por exemplo, na palestra em que proferiu na *Royal Society* em 27 de fevereiro de 1874.

Figura 3 – Um exemplar do Quincunx original e seu esquema.





Fonte: STILLMAN, Bruce, STEWART, David, WITKOWSKI (2010).

O quincunx permitiu que os estatísticos realizassem simulações da distribuição normal muito antes do advento dos computadores. O principal problema é a sua construção, pois a tarefa é difícil, mesmo para um bastante simples (PROSCHAN; ROSENTHAL, 2009, p. 11). O quincunx original (Figura 3, à esquerda) pode ser encontrado atualmente no Laboratório Galton em Londres. Nele pode-se ler uma inscrição feita pelo próprio Galton: instrumento para ilustrar o princípio da Lei dos Erros ou Dispersão. Francis Galton F.R.S. A sigla utilizada após o nome significa Membro da Sociedade Real (*Fellow of the Royal Society*).

CONCLUSÃO

A história quase nunca faz justiça aos descobridores originais. A curva normal é apenas mais um caso. Ela é conhecida como curva de Gauss que contribuiu apenas com uma aplicação e com isso, talvez, com sua popularização.

Com esse artigo pretendeu-se dar mais visibilidade a um dos principais modelos matemáticos que já foram criados. Agora em 2014, a curva normal ou mais apropriadamente distribuição ou modelo normal completa 280 anos, desde a sua criação por Abraão De Moivre em 1734. De Moivre derivou a curva, mas aparentemente ele não tinha noção de que ela poderia ser utilizada em conexão com dados experimentais. Ele era um analista puro e sem interesse em manipulação experimental de qualquer tipo. A utilização estatística da distribuição normal teve início com o trabalho de Laplace e de Gauss e que a derivaram independentemente e aparentemente sem o conhecimento do trabalho de De Moivre. As primeiras tabelas da integral do modelo foram publicadas, em 1799, em um livro sobre a refração da luz, pelo físico francês Christian Kramp (1760 – 1826) (DAVID, 2005).

O astrônomo alemão Friedrich Wilhelm Bessel (1784 – 1846) foi o primeiro a utilizar o termo erro provável, em 1815, em um artigo sobre a posição da estrela polar. Uma das derivações mais interessantes foi feita por outro astrônomo alemão John Herschel, em 1850, em que ele considerou uma distribuição de probabilidade bidimensional de erros ao mensurar a posição de uma estrela. Dez anos mais tarde, James C. Maxwell (1860) forneceu uma versão tridimensional desse mesmo argumento para encontrar a distribuição de probabilidade para a velocidade das moléculas de um gás, que ficou conhecida, entre os físicos, como "lei Maxwelliana da distribuição da velocidade" fundamental para a teoria cinética e a mecânica estatística (JAYNES, 2003, p. 703). Já a utilização do modelo para explicar dados que não fossem erros de observação deve-se ao astrônomo belga Adolphe Quetelet, que popularizou a ideia de que o método estatístico era uma disciplina fundamental e adaptável a qualquer campo de interesse no qual uma grande massa de dados fosse gerada. Embora Quetelet acreditasse que dados antropométricos, econômicos, da crimenologia, das ciências físicas, da botânica e da zoologia fossem todos fundamentalmente "normais", ele não foi o responsável pela criação do termo.

O modelo normal destaca-se não apenas pelas suas aplicações, mas também pelo grande número de cientistas que se envolveram na sua utilização, aperfeiçoamento e divulgação. Hoje ele é utilizado em praticamente todas as áreas em que se manipulam dados em quantidade, especialmente na área educacional quando se fazem as avaliações em larga escala. Os acertos em provas de algumas disciplinas, quando representadas graficamente, se assemelham ao modelo normal. Assim ele teve sua popularidade ampliada entre educadores e cientistas sociais. Contudo, nem sempre a sua utilização é feita de forma apropriada e quase nunca sua origem ou verdadeiro criador é lembrado.

REFERÊNCIAS

BOX, George Edward Pelham; MULLER, Mervin Edgar. A note on the generation of random normal deviates. **The Annals of Mathematical Statistical**. v. 29, n. 2, 1958, p. 610-11.

DAVID, Herbert A. Tables Related to the Normal Distribution: A Short History. **The American Statistician**. v. 59, n. 4, Nov. 2005, p. 309-11.

DAW, R. H. Why the Normal Distribution. **Journal of the Institute of Actuaries Students' Society [JSS]**. v. 18, n. 1, 1966, p. 2-15.

DAW, R. H.; PEARSON, Egon Sharpe. Studies in the history of probability and statistics: Abraham de Moivre's 1733 derivation of the normal curve: a bibliographical note. **Biometrika**, v. 59, n. 3, 1972, p. 677-80.

EDGEWORTH, Francis Ysidro. The Generalised Law of Error, or Law of Great Numbers. **Journal of the Royal Statistical Society**, v. 69, n. 3, Sep., 1906, p. 497-539.

EISENHART, Churchill; BIRNBAUM, Allan. Anniversaries in 1966-67 of Interest to Statisticians: Part II: Tercentennials of Arbuthnot and De Moivre. **The American Statistician**, v. 21, n. 3, Jun., 1967, p. 22-9.

FENDLER, Lynn; MUZAFFAR, Irfan. The History of The Bell Curve: Sorting and The Idea of Normal. **Educational Theory.** v. 58, n. 1, Feb 2008, p. 63-82.

JAYNES, E. T. **Probability Theory:** the Logic of Science. Cambridge (MA): the Cambridge University Press. 2003. 724 p.

KUNERT, Joachim; MONTAG, Astrid; PÖHLMANN, Sigrid. The quincunx: history and mathematics. **Springer: Statistical Papers**. v. 42, 2001, p. 143-69.

MAXWELL, James Clerk. Illustrations of the dynamical theory of gases. 1860. In: NIVEN, W. D (Ed.). **The Scientific Papers of James Clerk Maxwell**. vol. 1, Paris: Dover Phoenix Edition, 2003, chapter 10, p. 377-410.

METROPOLIS, N. C. The beginning of the Monte Carlo Method. **Los Alamos Science special issue**. 1987, p. 125–30.

BRASIL. Ministério da Educação. Secretaria da Educação Média e Tecnológica. **Parâmetros Curriculares Nacionais (PCN+) - Ciências da Natureza e suas Tecnologias.** Brasília: MEC, 2002.

PEARSON, Karl. Notes on the History of Correlation. **Biometrika**. v. 13, n.1 , October 1920, p. 25-45.

PEARSON, Karl. Historical note on the origin of normal curve of errors. **Biometrika**, v. 16, n. 3/4, Dec. 1924, p. 402-404.

PROSCHAN, Michael A.; ROSENTHAL, Jeffrey S. Beyond the Quintessential Quincunx. **Department of Statistics. University of Toronto**. Technical Report No. 0905 September 8, 2009. 12 p.

STAHL, Saul. The evolution of normal distribution. **Mathematics Magazine**. v. 79, n. 2, April 2006. p. 96-112.

STIGLER, Stephen. M. **The History of Statistics: The measurement of Uncertainty before 1900**. Cambridge (MA): Belkarnap Press of Harvard University Press, 2000. 410 p.

STIGLER, Stephen. M. **Statistics on the Table: the history of statistical concepts and methods**. Cambridge (MA): Belkarnap Press of Harvard University Press, 2002. 488 p.

STILLMAN, Bruce; STEWART, David; WITKOWSKI (Eds). **Evolution: the Molecular Landscape**. Cold Spring Harbor Symposia on Quantitative Biology LXXIV. Cold Spring Harbor Laboratory Press, 2010.

THIELE, T. N. Laws of Errors. **The Annals of Mathematical Statistics**, v. 2, n. 2, May, 1931, p. 169-72.

TODHUNTER, Isaac. **A History of the Mathematical Theory of Probability:** From the Time of Pascal to That of Laplace. London: Elibron Classics, 2005.

YOUDEN, William John. **Experimentation and Measurement**. NIST Special Publication 672. U.S Departament of Commerce. Calibration Program. Statistical Engineering Division. 1997, 128 p.

WALKER, Helen M. Bi-Centenary of the Normal Curve. **Journal of the American Statistical Association**. v. 29, n. 185, March 1934, p. 72-75.

(*) Figuras de domínio público disponíveis em vários sites.

RECEBIDO EM: 01.02.2014.

CONCLUÍDO EM: 01.04.2014.