6

Variáveis Aleatórias Contínuas

6.1 CARACTERIZAÇÃO DE UMA VARIÁVEL ALEATÓRIA CONTÍNUA

Muitas variáveis aleatórias que surgem na vida de um engenheiro ou de um profissional da informática têm natureza eminentemente contínua, tais como:

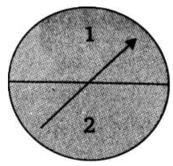
- tempo de resposta de um sistema computacional;
- rendimento de um processo químico;
- tempo de vida de um componente eletrônico;
- resistência de um material etc.

Outras vezes, há variáveis aleatórias discretas, com grande número de possíveis resultados, em que é preferível usar um modelo aproximado contínuo no lugar do modelo exato discreto. É o caso de:

- número de transações por segundo de uma CPU;
- número de defeitos numa amostra de 5.000 itens etc.

Para entender as peculiaridades das variáveis aleatórias contínuas, imagine o seguinte experimento.

Exemplo 6.1a Um círculo é dividido em dois setores de mesmo tamanho (180° cada um), aos quais são atribuídos os números 1 e 2. Um ponteiro é preso ao centro do círculo e girado, conforme mostra a figura ao lado. Seja a variável aleatória discreta X = *número do setor apontado quando o ponteiro pára de girar*. A distribuição de probabilidades de X, considerando que todos os pontos sejam equiprováveis, pode ser especificada pela função de probabilidade da Figura 6.1.



x	$p(x)$
1	½
2	½
Total	1

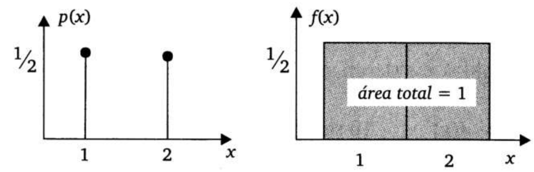


Figura 6.1 *Três formas de apresentação da função de probabilidade do experimento aleatório do exemplo 6.1a.*

A representação com gráfico de hastes é típica para variáveis discretas. Apresentamos, também, um gráfico em forma de histograma, em que as probabilidades podem ser representadas por área. No caso do exemplo em questão, as bases dos retângulos são iguais à unidade, o que faz com que a área seja igual à altura do retângulo.[1]

Exemplo 6.1b Considere, agora, o círculo dividido em quatro setores de mesmo tamanho (90° cada um). A distribuição de probabilidades da variável aleatória discreta X = *número do setor apontado quando o ponteiro pára de girar* é apresentada na Figura 6.2.

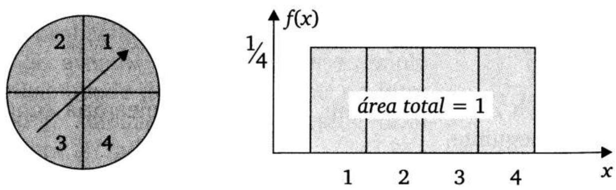


Figura 6.2 *Representação gráfica da função de probabilidade do experimento aleatório do Exemplo 6.1b.*

1 Ao representarmos probabilidades por áreas, devemos tomar o cuidado para que área total seja igual à unidade.

Exemplo 6.1c Imagine que o círculo seja dividido em 8, 16 e 32 setores. A Figura 6.3 mostra a função de probabilidade de *X* em cada caso.

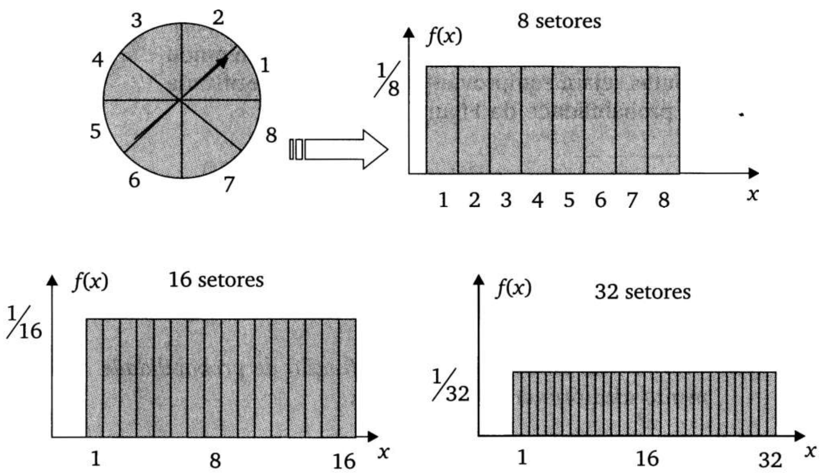


Figura 6.3 *Representações gráficas das funções de probabilidade dos experimentos aleatórios do Exemplo 6.1c.*

É fácil verificar que à medida que aumentamos o número de divisões no círculo, o número de possíveis setores (resultados de uma variável aleatória discreta) vai aumentando, e a probabilidade de cada resultado ocorrer (representada pela área de um retângulo) vai sendo reduzida. Teoricamente, o círculo pode ser dividido em infinitos setores, o que torna inviável a representação tabular ou gráfica da distribuição de probabilidades, da forma como fizemos no Exemplo 6.1. Em termos matemáticos, teríamos:

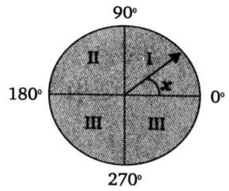
$$P(x) = P(X = x) = \lim_{n \to \infty} \frac{1}{n} = 0, \quad \forall x = 1, 2, \ldots \tag{6.1}$$

Uma alternativa melhor é definir uma variável aleatória contínua, como veremos na seção seguinte.

6.1.1 Função densidade de probabilidade

Exemplo 6.2 Considere um círculo, com medidas de ângulos, em graus, a partir de determinada origem, como mostra a figura ao lado. Nesse círculo, há um ponteiro que é colocado a girar.

Seja a variável aleatória contínua X = *ângulo formado entre a posição que o ponteiro pára e a linha horizontal do lado direito.* Considerando que não existe região de preferência para o ponteiro parar, a distribuição de probabilidade de X pode ser representada por uma função que assume um valor constante e positivo em todo o intervalo $[0^o, 360°)$, de tal forma que as probabilidades possam ser representadas por áreas sob a curva dessa função. Como certamente vai ocorrer um resultado em $[0^o, 360°)$, então a área sob a função neste intervalo deve ser igual a 1, e nula fora deste intervalo. A Figura 6.4 ilustra a distribuição de probabilidades de X, através da chamada *função densidade de probabilidade,* e mostra a relação entre uma área e um evento.



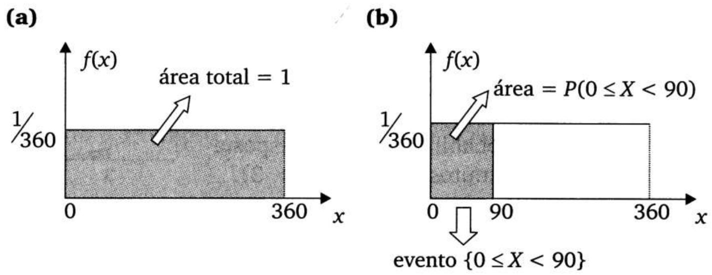
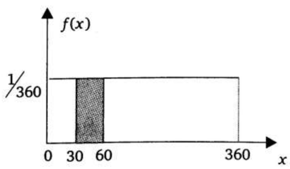


Figura 6.4 *(a) Função densidade de probabilidade da variável aleatória do Exemplo 6.2; e (b) Probabilidade do evento $\{0 \leq X < 90\}$, representada por uma área.*

Note que os eventos associados a uma variável aleatória contínua são intervalos (ou coleção de intervalos) dos números reais. Com base na função densidade de probabilidade, podemos calcular probabilidades de eventos desse tipo. Por exemplo, qual é a probabilidade do ponteiro parar no intervalo $[30°, 60°]$? Tomando a área do retângulo indicado na figura ao lado, temos:



$$P(30 \leq X \leq 60) = \frac{1}{360} \cdot (60 - 30) = \frac{30}{360} = \frac{1}{12}$$

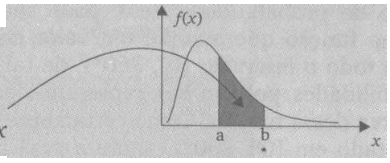
Observe que a inclusão ou exclusão dos extremos não altera a probabilidade, pois uma linha tem área nula. Ou seja, para uma variável aleatória contínua, a probabilidade de ocorrer um particular valor é igual a zero.

As probabilidades de eventos associados a uma variável aleatória contínua X podem ser calculadas através de uma **função densidade de probabilidade f** que deve satisfazer:

a) $f(x) \geq 0, \ \forall x \in \Re$ e

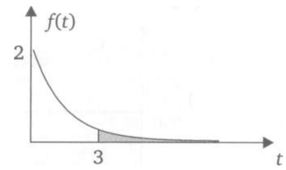
b) $\int_{-\infty}^{+\infty} f(x)\,dx = 1$

Se $A = [a, b]$, então $P(A) = \int_a^b f(x)\,dx$



Exemplo 6.3 Seja a variável aleatória T definida como o tempo de resposta na consulta a um banco de dados, em minutos. Suponha que essa variável aleatória tenha a seguinte função densidade de probabilidade:

$$f(t) = \begin{cases} 2e^{-2t}, & \text{para } t \geq 0 \\ 0, & \text{para } t < 0 \end{cases}$$



Vamos calcular a probabilidade de a resposta demorar mais do que 3 minutos, isto é, $P(T > 3)$.

$$P(T > 3) = \int_3^{+\infty} f(t)dt = \int_3^{+\infty} 2e^{-2t}dt = 2\left[-\frac{1}{2}e^{-2t}\right]_3^{+\infty} = 0 + e^{-2(3)} = e^{-6}$$

6.1.2 Função de distribuição acumulada

Como X é uma variável aleatória contínua com função de densidade de probabilidade f, definimos sua *função de distribuição acumulada* por:

$$\boxed{F(x) = P(X \leq x) = \int_{-\infty}^{x} f(s)ds, \quad \forall x \in \Re} \tag{6.2}$$

Considere a função densidade de probabilidade do Exemplo 6.3:

$$f(t) = \begin{cases} 2e^{-2t}, & \text{para } t \geq 0 \\ 0, & \text{para } t < 0 \end{cases}$$

Vamos obter a função de distribuição acumulada. Como a expressão matemática se altera no ponto zero, devemos considerar os dois seguintes casos:

para $t < 0$,

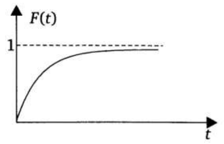
$$F(t) = \int_{-\infty}^{t} f(s)ds = \int_{-\infty}^{t} 0ds = 0$$

e para $t \geq 0$,

$$F(t) = \int_{-\infty}^{t} f(s)ds = \int_{-\infty}^{0} 0ds + \int_{0}^{t} 2e^{-2s}ds = 0 + \left[-e^{-2s}\right]_{0}^{t} = 1 - e^{-2t}$$

Resumindo, a função de distribuição acumulada da variável aleatória *T é* dada por:

$$F(t) = \begin{cases} 1 - e^{-2t}, & \text{para } t \geq 0 \\ 0, & \text{para } t < 0 \end{cases}$$



Cabe observar que é possível obter qualquer probabilidade através da função de distribuição acumulada. Para $a < b$, temos:

$$P(X < a) = P(X \leq a) = F(a) \tag{6.3}$$

$$P(X > b) = 1 - F(b) \tag{6.4}$$

$$P(a < X < b) = F(b) - F(a) \tag{6.5}$$

Retomando o Exemplo 6.3, o cálculo de $P(T > 3)$ pode ser feito aplicando (6.4):

$$P(T > 3) = 1 - P(T \leq 3) = 1 - F(3) = 1 - [1 - e^{-2(3)}] = e^{-6}$$

Dada a função de distribuição acumulada F, podemos obter a função densidade de probabilidade f por:

$$f(x) = \frac{d}{dx}F(x) \tag{6.6}$$

para todo ponto x em que *F é* derivável.[2] Assim, a função *F* também caracteriza a distribuição de probabilidades de uma variável aleatória.

6.1.3 Valor esperado e variância

Uma variável aleatória contínua *X,* com função densidade de probabilidade f, tem *valor esperado* e *variância* definidos por:

$$\boxed{\begin{aligned} \mu &= E(X) = \int_{-\infty}^{+\infty} xf(x)dx \\ \sigma^2 &= V(X) = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx \end{aligned}}$$

(6.7)

(6.8)

2 No conjunto finito de pontos em que *F* não é derivável, podemos arbitrar valores para/.

As interpretações dessas medidas podem ser feitas de forma análoga ao caso discreto. Além disso, todas as propriedades enunciadas para o caso discreto continuam válidas para o caso contínuo, em especial

$$\boxed{V(X) = E(X^2) - \mu^2} \qquad (6.9)$$

onde: $E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x)dx$

Retomando o Exemplo 6.3, em que a variável aleatória T era caracterizada por

$$f(t) = \begin{cases} 2e^{-2t}, & \text{para } t \geq 0 \\ 0, & \text{para } t < 0 \end{cases}$$

temos: $\mu = E(T) = \int_{-\infty}^{+\infty} tf(t)dt = \int_{-\infty}^{0} t0dt + \int_{0}^{+\infty} t2e^{-2t}dt = 0 + 2\int_{0}^{+\infty} te^{-2t}dt$

Integrando $\int_{0}^{+\infty} t2e^{-2t}dt$ por partes, obtemos $\mu = ½$.

Temos, também,

$$E(T^2) = \int_{-\infty}^{+\infty} t^2 f(t)dt = \int_{-\infty}^{0} t^2 0dt + \int_{0}^{+\infty} t^2 2e^{-2t}dt = 0 + 2\int_{0}^{+\infty} t^2 2e^{-2t}dt$$

Com certo esforço matemático, obtemos $E(T^2) = ½$. Então:

$$\sigma^2 = V(T) = E(T^2) - \mu^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4}$$

EXERCÍCIOS

1. Seja um ponto escolhido aleatoriamente no intervalo [0, 1].
   a) Apresente uma função densidade de probabilidade para este experimento.
   b) Obtenha a função de distribuição acumulada.
   c) Calcule o valor esperado e a variância.

2. Um profissional de Computação observou que seu sistema gasta entre 20 e 24 segundos para realizar determinada tarefa. Considere a probabilidade uniforme em [20, 24], isto é, todo subintervalo de mesma amplitude em

[20, 24] tem a mesma probabilidade. Como pode ser descrita, gráfica e algebricamente, a função densidade de probabilidade? Sob essa densidade, calcule:

a) $P(X > 23)$;

b) $E(X)$;

c) $V(X)$.

3. Com respeito ao exercício anterior, mas supondo probabilidades maiores em torno de 22 segundos e a densidade decrescendo, simétrica e linearmente, até os extremos 2C e 24 segundos. Como pode ser descrita, gráfica e algebricamente, a função densidade de probabilidade? Sob essa densidade, calcule:

a) $P(X > 23)$;

b) $E(X)$;

c) $V(X)$.

Comparando os gráficos das funções de densidade de probabilidade dos dois exercícios, você acha razoáveis as diferenças encontradas nos três itens?

4. Seja X uma variável aleatória com função de distribuição acumulada

$$F(x) = \begin{cases} 1 - e^{-x}, & \text{para } x \geq 0 \\ 0, & \text{para } z < 0 \end{cases}$$

Obtenha a função densidade de probabilidade de X.

5. Seja X com função densidade de probabilidade dada por

$$f(x) = \begin{cases} x, & \text{para } 0 \leq x < 1 \\ 2 - x, & \text{para } 1 \leq x < 2 \\ 0, & \text{para } x \notin [0, 2) \end{cases}$$

Calcule:

a) $P(0 < X < 5)$ **b)** $P(0 < X < 1)$ **c)** $P(1/3 < X < 3/2)$

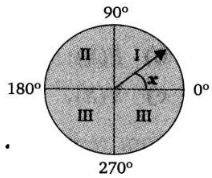
d) $E(X)$ **e)** $V(X)$

6.2 PRINCIPAIS MODELOS CONTÍNUOS

Nesta seção serão descritos três modelos contínuos bastante conhecidos.

6.2.1 Distribuição uniforme

Relembremos o Exemplo 6.2, onde tínhamos um círculo e um ponteiro que era colocado a girar. A variável aleatória de interesse era X = *ângulo formado entre a posição que o ponteiro pára e a linha horizontal do lado direito.* Supôs-se, também, não existir região de preferência para o ponteiro parar. Nessas codições, podemos considerar que todo intervalo de mesma amplitude, contido em $[0^o, 360°)$, tem a mesma probabilidade de ocorrência. É um experimento típico em que a chamada *distribuição uniforme* é apropriada.



Uma variável aleatória X tem *distribuição uniforme* de parâmetros α e β, sendo β > α, se sua densidade é especificada por:

$$\boxed{f(x) = \begin{cases} \dfrac{1}{\beta - \alpha}, & \text{para } x \in [\alpha, \beta] \\ 0, & \text{para } x \notin [\alpha, \beta] \end{cases}} \qquad \textbf{(6.10)}$$

Em conseqüência, sua distribuição acumulada é dada por (ver Figura 6.5):

$$F(x) = \begin{cases} 0, & \text{para } x < \alpha \\ \dfrac{x - \alpha}{\beta - \alpha} & \text{para } \alpha \le x < \beta \\ 1, & \text{para } x \ge \beta \end{cases} \qquad \textbf{(6.11)}$$

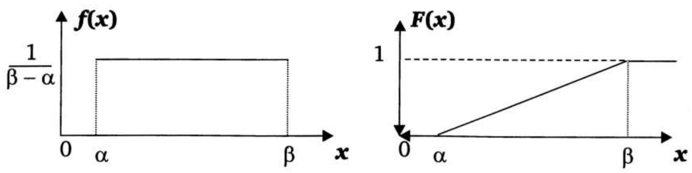


Figura 6.5 *Representação gráfica da função densidade de probabilidade e da função de distribuição acumulada de uma variável aleatória com distribuição uniforme em [α, β].*

O valor esperado e a variância de uma distribuição uniforme são:

$$E(X) = \frac{\alpha + \beta}{2} \qquad \textbf{(6.12)}$$

$$V(X) = \frac{(\beta - \alpha)^2}{12} \qquad (6.13)$$

Note que o valor esperado da distribuição uniforme é exatamente o ponto médio do intervalo [α, β], ou seja, nessa distribuição fica evidente que μ representa o centro de gravidade da massa descrita pela função densidade de probabilidade.

6.2.2 Distribuição exponencial

O modelo exponencial tem forte relação com o modelo discreto de Poisson. Enquanto a distribuição de Poisson pode ser usada para modelar o *número de ocorrências* em um período contínuo (de tempo ou de comprimento), a distribuição exponencial pode modelar a variável aleatória contínua que representa o *intervalo* (de tempo ou de comprimento) entre as ocorrências. Exemplos:

a) tempo (em minutos) até a próxima consulta a uma base de dados;

b) tempo (em segundos) entre pedidos a um servidor;

c) distância (em metros) entre defeitos de uma fita.

A distribuição exponencial pode ser usada quando as suposições de Poisson (independência entre as ocorrências e taxa média de ocorrência constante no intervalo considerado) estiverem satisfeitas. A Figura 6.6 ilustra a relação entre as duas distribuições.

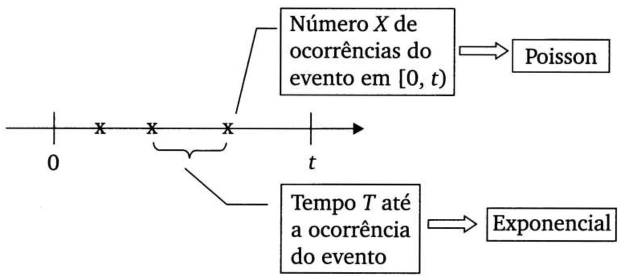
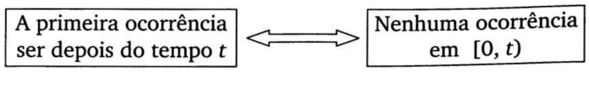


Figura 6.6 *Relação entre a distribuição de Poisson e a exponencial*

Para chegarmos à formulação matemática da distribuição exponencial, vamos considerar a equivalência entre os dois seguintes eventos:



Sejam as variáveis aleatórias:

X_t = número de ocorrências no intervalo de tempo [0, t); e
T = tempo entre as ocorrências.

Sendo λ a taxa média de ocorrências por unidade de tempo, então, considerando independência entre as ocorrências, X_t tem distribuição de Poisson com parâmetro λt. E a equivalência entre os dois eventos pode ser expressa por:

$$\boxed{T > t} \Longleftrightarrow \boxed{X_t = 0}$$

Logo,

$$P(T > t) = P(X_t = 0) = \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = e^{-\lambda t}$$

aplicação da expressão de Poisson

Usando o evento complementar, podemos definir para todo $t > 0$ a função de distribuição acumulada de uma variável aleatória T com distribuição exponencial:

$$\boxed{F(t) = P(T \le t) = 1 - e^{-\lambda t}} \tag{6.14}$$

Em conseqüência, para $t > 0$ temos a função densidade de probabilidade dada por:

$$\boxed{f(t) = \frac{d}{dt} F(t) = \lambda e^{-\lambda t}} \tag{6.15}$$

Para $t \le 0$, definimos $F(t) = f(t) = 0$ (ver a Figura 6.7).

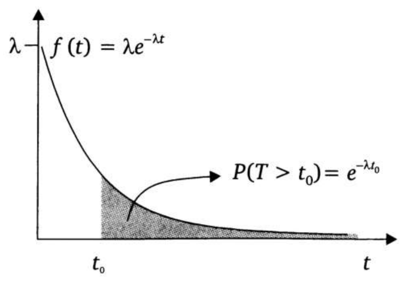


Figura 6.7 *Representação gráfica da função densidade de probabilidade de uma variável aleatória com distribuição exponencial.*

Em geral, é mais fácil partir do complemento de F(t) para calcular as probabilidades, ou seja, para $t > 0$,

$$\boxed{P(T > t) = e^{-\lambda t}} \qquad (6.16)$$

Exemplo 6.3 (continuação) Dada a variável aleatória T = *tempo de resposta na consulta a um banco de dados por minutos)* com função densidade de probabilidade

$$f(t) = \begin{cases} 2e^{-2t} & \text{para } t \geq 0 \\ 0, & \text{para } t < 0 \end{cases}$$

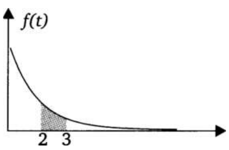
ou seja, uma exponencial com $\lambda = 2$, calcular a probabilidade da consulta demorar mais que 3 minutos, isto é, $P(T > 3)$. Podemos partir da função de densidade, fazendo:

$$P(T > 3) = \int_3^{+\infty} f(t)dt = \int_3^{+\infty} 2e^{-2t}dt = e^{-6}$$

Ou podemos usar (6.16), obtendo:

$$P(T > 3) = e^{-2(3)} = e^{-6}$$

Considere, agora, o cálculo da probabilidade $P(2 \leq T \leq 3)$, isto é, a próxima consulta ocorrer no intervalo de 2 a 3 minutos. Podemos fazer



$$P(2 \leq T \leq 3) = \int_2^3 2e^{-2t}dt$$

ou usar (6.5):

$$P(2 \leq T \leq 3) = P(T \geq 2) - P(T > 3) = e^{-2(2)} - e^{-2(3)} = e^{-4} - e^{-6} =$$
$$= 0{,}0158$$

Para uma variável aleatória T, com distribuição exponencial de parâmetro λ, temos:

$$\boxed{E(T) = \frac{1}{\lambda}} \qquad (6.17)$$

$$\boxed{V(T) = \frac{1}{\lambda^2}} \qquad \textbf{(6.18)}$$

Um exemplo do cálculo do valor esperado e da variância de uma exponencial foi feito na Seção 6.1.3. Observe que podemos obter os mesmos resultados com (6.17) e (6.18).

EXERCÍCIOS

6. O tempo de vida (em horas) de um transistor é uma variável aleatória T com distribuição exponencial. O tempo médio de vida do transistor é de 500 horas.

a) Calcule a probabilidade de o transistor durar mais do que 500 horas.

b) Calcule a probabilidade de o transistor durar entre 300 e 1000 horas.

c) Sabendo-se que o transistor já durou 500 horas, calcule a probabilidade de ele durar mais 500 horas.

7. Usando a expressão de probabilidade condicional (Capítulo 4), mostrar que para $s, t > 0$, vale a seguinte relação para uma variável aleatória T exponencial:

$$P(T > s + t \mid T > s) = P(T > t)$$

Essa propriedade é conhecida como "falta de memória", pois não importa o que aconteceu no passado ($T \leq s$), mas apenas a partir do momento em que se inicia a observação, que pode ser considerado como o instante zero. Nesse contexto, a distribuição exponencial é inadequada para representar "tempo de vida" de itens que sofrem efeito de fadiga.

6.2.3 Distribuição normal

A normal é considerada a distribuição de probabilidades mais importante, pois permite modelar uma infinidade de fenômenos naturais e, além disso, possibilita realizar aproximações para calcular probabilidades de muitas variáveis aleatórias que têm outras distribuições. É muito importante também na inferência estatística, como será observado nos capítulos seguintes.

A distribuição normal é caracterizada por uma função de probabilidade, cujo gráfico descreve uma curva em forma de sino, como mostra a Figura 6.8. Essa forma de distribuição evidencia que há maior probabilidade de a variável aleatória assumir valores próximos do centro.

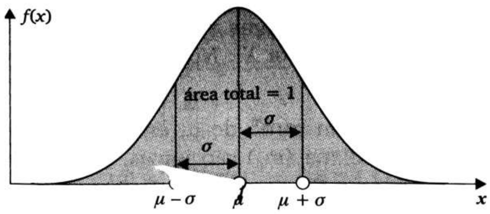


Figura 6.8 *Representação gráfica da função densidade de probabilidade normal e a indicação de seus dois parâmetros:* μ e σ.

Dados os parâmetros μ e R e σ > 0, a função densidade de probabilidade da normal é dada por:

$$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \ -\infty < x < +\infty}$$

Com certo esforço matemático, é possível mostrar que o *valor esperado* *e a variância* da dis

$$\boxed{\begin{array}{c} E(X) = \mu \\ V(X) = \sigma^2 \end{array}}$$

A Figura 6.9 mostra diferentes curvas normais, em função dos valores de e a. As distribuições da Figura 6.8 podem representar, por exemplo, medidas da dureza de aço produzido sob diferentes condições. A distribuição (1) representa a dureza do aço em uma situação padrão; e a distribuição (2), as medidas de dureza do aço após um processo de melhoria da qualidade, em que aumentou a dureza média. A distribuição (3) representa as medidas de dureza do aço quando o processo está sob rígido controle; enquanto a distribuição (4) quando fora de controle, o que acarreta aumento na variabilidade.

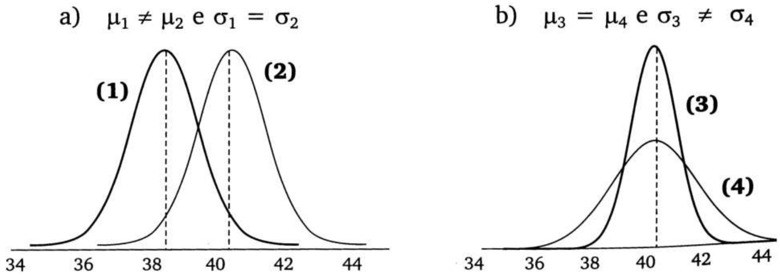


Figura 6.9 *Diferentes distribuições normais em função dos parâmetros* μ *e* σ

Na seqüência, representaremos uma variável aleatória X com distribuição normal de média μ e variância σ^2 por $X : N(\mu, \sigma^2)$. Seguem outras características do modelo normal:

- a curva é simétrica em torno de μ, em conseqüência, os valores da média (μ) e da mediana (m_d) são iguais, e também $P(X < \mu - \alpha) = P(X > \mu + \alpha)$, $\forall\ \alpha \in \Re$;
- teoricamente, a curva prolonga-se de $-\infty$ a $+\infty$, sendo $\lim_{x \to \pm\infty} f(x) = 0$;
- a área total sob a curva é igual a 1, ou seja, $\int_{-\infty}^{+\infty} f(x)dx = 1$;
- qualquer combinação linear de variáveis aleatórias normais é também uma variável aleatória normal; em especial, se X_1 e X_2 são variáveis aleatórias independentes e $X_1 : N(\mu_1, \sigma_1^2)$ e $X_2 : N(\mu_1, \sigma_2^2)$, então $\forall \alpha$, $b \in \Re$, $Y = aX_1 + bX_2$ tem distribuição normal com

$$E(Y) = a\mu_1 + b\mu_2 \tag{6.22}$$

$$V(Y) = a^2\, \sigma_1^2 + b^2\, \sigma_2^2 \tag{6.23}$$

- afastamentos da média, em unidades de desvio padrão, preservam a mesma área sob a curva, independentemente dos valores de μ e σ (ver a Figura 6.10).

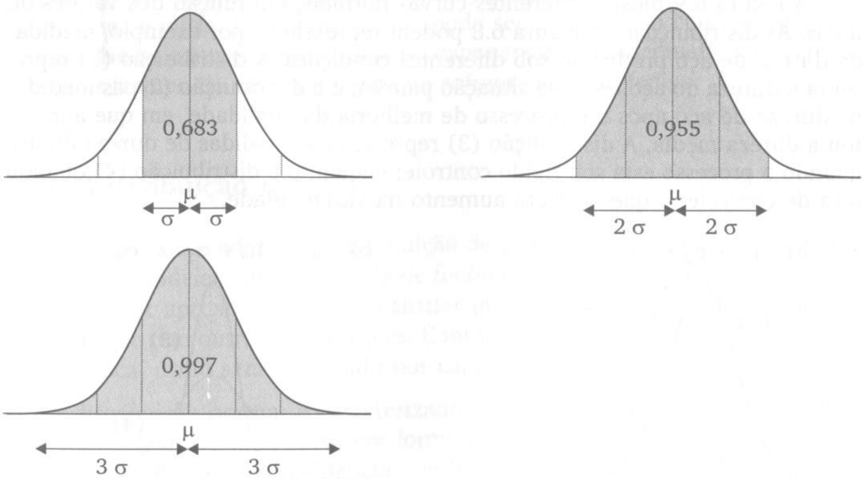


Figura 6.10 *Afastamentos da média, em unidades de desvio padrão, preservam a mesma área sob a curva normal.*

Seja $X : N(\mu, \sigma^2)$, então a variável aleatória

$$\boxed{Z = \frac{X - \mu}{\sigma}} \qquad \textbf{(6.24)}$$

tem distribuição normal com média zero e desvio padrão unitário, ou seja, Z : N(0,1), que também é conhecida como **distribuição formal padrão.** Qualquer área (probabilidade) sob a densidade de *X* pode ser avaliada sob a densidade de Z, conforme ilustra a Figura 6.11. Dessa forma, qualquer problema relativo a uma distribuição normal pode ser pensado em termos da distribuição normal padrão.

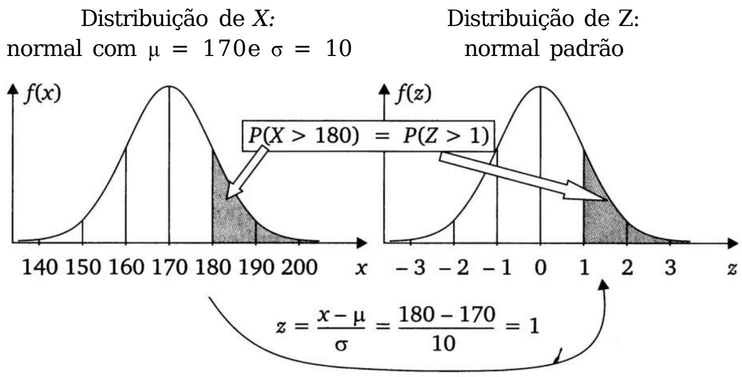


Figura 6.11 *Transformação do evento {X > 180}, da distribuição normal de parâmetros μ = 170 e σ = 10, em um evento da distribuição normal padrão: {Z > 1}.*

Tabela da distribuição normal padrão

Como vimos, as probabilidades de uma variável com distribuição normal podem ser representadas por áreas sob a curva da distribuição normal padrão. No apêndice, apresentamos a Tabela 3, que relaciona valores positivos de com áreas sob a cauda superior da curva. Os valores de *z* são apresentados com duas decimais. A primeira decimal fica na coluna da esquerda e a segunda decimal na linha do topo da tabela. A Figura 6.12 mostra como podemos usar a Tabela 3 para encontrar uma área sob a cauda superior da curva.

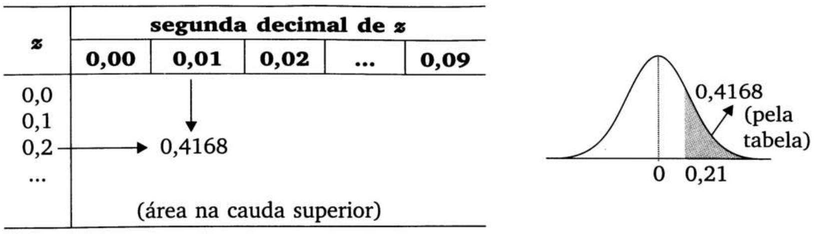
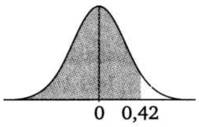


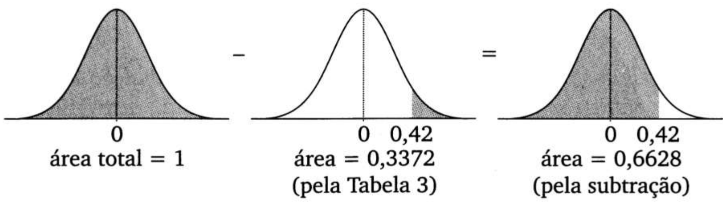
Figura 6.12 *Ilustração do uso da tabela da distribuição normal padrão (Tabela 3 do apêndice) para encontrar P(Z > 0,21).*

A área 0,4168 corresponde à probabilidade P(Z > 0,21) = 1 - Φ (0,21), onde Φ representa a *função de distribuição acumulada da normal padrão.* Ou seja, a Tabela 3 fornece os valores 1 - Φ(z), para z = 0,01, 0,02, ..., 3,00.

Exemplo 6.4 Seja Z uma variável aleatória com distribuição normal padrão. Vamos usar a Tabela 3 para encontrar as seguintes probabilidades:

a) P(Z < 0,42). Esta probabilidade corresponde à área da distribuição normal padrão indicada ao lado. Podemos obter esta área, fazendo a seguinte operação:

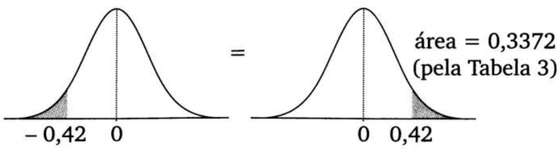




Mais formalmente,

$$\Phi(0{,}42) = P(Z < 0{,}42) = 1 - P(Z > 0{,}42) = 1 - 0{,}3372 = 0{,}6628$$

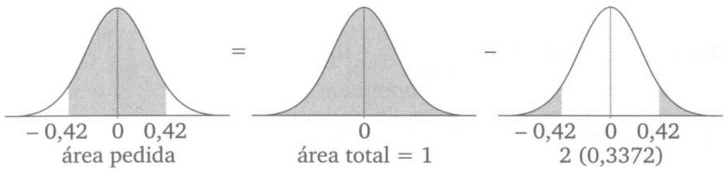
b) P(Z < 0,42). O esquema seguinte mostra esta probabilidade em termos de área e como podemos usar a simetria da curva para obtê-la na Tabela 3.



Ou seja,

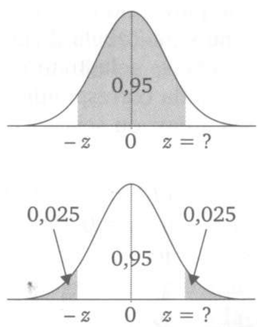
$P(Z < -0{,}42) = P(Z > 0{,}41) = 0{,}3372$

c) $P(-0{,}42 < Z < 0{,}42)$.



Então, $P(-0{,}42 < Z < 0{,}42) = 1 - 2\,(0{,}3372) = 0{,}3256$.

Exemplo 6.5 Na distribuição normal padrão, qual é o valor de z, tal que $P(-z < Z < z) = 0{,}95$? (Veja figura ao lado.)



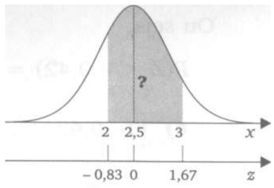
Considerando a simetria da curva normal e o fato de a área total sob a curva ser igual a 1 (um), podemos transformar esta pergunta em: *Qual é o valor de z tal que $P(Z > z) = 0{,}025$?* A figura ao lado ilustra a equivalência entre as duas perguntas.

Entrando com o valor de área 0,025 na Tabela 3 do apêndice, encontramos o valor de z igual a 1,96. Esse processo é ilustrado a seguir.

z	0,00	0,01	...	0,06	...	0,09
...						
1,9				0,025		
...						

Exemplo 6.6 Suponha que a absorção de água (%) em certo tipo de piso cerâmico tenha distribuição normal com média 2,5 e desvio padrão 0,6. Selecionando, aleatoriamente, uma unidade desse piso, qual é a probabilidade de ele acusar absorção de água entre 2% e 3,5%?

Solução: Primeiramente, precisamos transformar os valores de absorção de água (x) em valores padronizados (z), por (6.24), isto é,

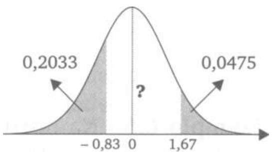


$$z = \frac{x - \mu}{\sigma} = \frac{x - 2,5}{0,6}$$

Para x = 2, temos: $z = \frac{2 - 2,5}{0,6} = -0,83$

e para x = 3,5, temos: $z = \frac{3,5 - 2,5}{0,6} = 1,67.$

Usando a Tabela 3 do apêndice, encontramos para 2 = 0,83 *e z* = 1,67 as respectivas áreas nas extremidades da curva: 0,2033 e 0,0475 (lembrando que para valores negativos de *z*, como - 0,83, procuramos na Tabela 3 seu valor absoluto, 0,83). É fácil observar, pela figura ao lado, que a probabilidade desejada corresponde ao complemento da soma dessas áreas, ou seja:



$$P(2 < X < 3,5) = P(-0,83 < Z < 1,67) = 1 - (0,2033 + 0,0475) =$$
$$= 0,7492.$$

EXERCÍCIOS

8. Seja Z uma variável aleatória com distribuição normal padrão. Calcule:

a) $P(Z > 1,65)$;

b) $P(Z < 1,65)$;

c) $P(-1 < Z < 1)$;

d) $P(-2 < Z < 2)$;

e) $P(-3 < Z < 3)$;

f) $P(Z > 6)$;

g) o valor de z, tal que $P(-z < Z < z) = 0,90$;

h) o valor de z, tal que $P(-z < Z < z) = 0,99$.

9. Suponha que o tempo de resposta na execução de um algoritmo é uma variável aleatória com distribuição normal de média 23 segundos e desvio padrão de 4 segundos. Calcule:

a) a probabilidade de o tempo de resposta ser menor do que 25 segundos;

b) a probabilidade de o tempo de resposta ficar entre 00 e 30 segundos.

10. Certo tipo de conserva tem peso líquido (X_x) com média de 900 g e desvio padrão de 10 g. A embalagem tem peso (X_2) com média de 100 g e desvio padrão de 4 g. Suponha X_1 e X_2 independentes e com distribuições norfriais.

a) Qual é a probabilidade de o peso bruto ser superior a 1.020 g?

b) Qual é a probabilidade do peso bruto estar entre 980 e 1.020 g?

6.3 A NORMAL COMO LIMITE DE OUTRAS DISTRIBUIÇÕES

Muitas distribuições de probabilidade aproximam-se da distribuição normal. É o caso da binomial quando *n* é grande e da Poisson quando λ é grande.

6.3.1 Aproximação normal à binomial

Nos experimentos binomiais, quando *n é* muito grande, o uso da função de probabilidade binomial é impraticável, pois os coeficientes binomiais tornam-se exageradamente grandes. Já vimos que nos casos em que *n* é grande e *p é* muito pequeno, podemos usar a distribuição de Poisson para calcular, aproximadamente, as probabilidades de uma binomial. Quando *n é* grande e p não é próximo de 0 ou de 1, a distribuição normal pode ser usada para calcular, aproximadamente, as probabilidades de uma binomial.

A Figura 6.13 apresenta gráficos das distribuições de probabilidades binomiais com n = 1, 10 e 50 e p = 0,5 e 0,2.

Observando a Figura 6.13, verificamos que quando $n = 50$, a forma da distribuição binomial é parecida com a curva de uma distribuição normal. Observe, ainda, que se $p = 0,5$, a aproximação já parece razoável para $n = 10$.

De maneira geral, as condições para fazer uma aproximação da distribuição binomial para a normal são:

1) n grande e

2) *p* **não** muito próximo de 0 (zero) ou de **1** (um).

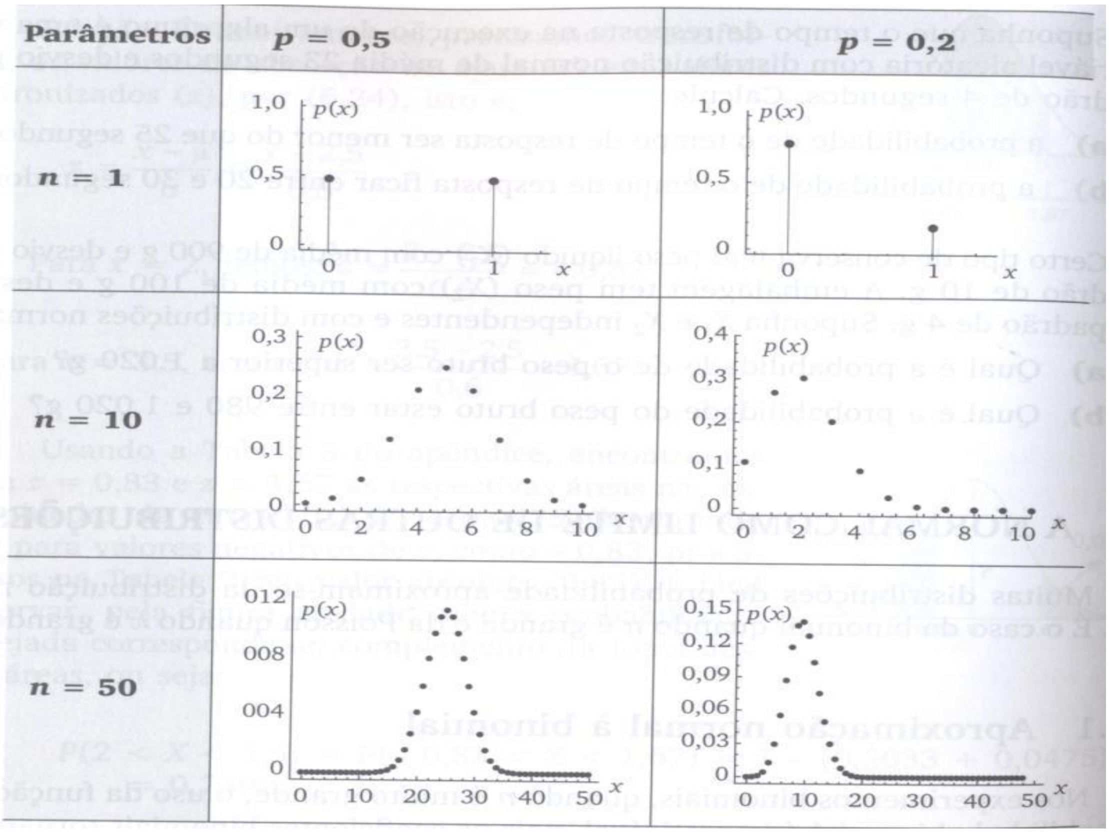


Figura 6.13 *Distribuições binomiais para diferentes valores de* n *e* p.

Uma regra prática, sugerida por vários autores, considera a aproximação razoável se as duas seguintes inequações estiverem satisfeitas:

$$np \geq 5 \text{ e} \qquad (6.25)$$

$$n(1 - p) \geq 5 \qquad (6.26)$$

Os parâmetros μ e σ da distribuição normal devem-se identificar ao valor esperado e ao desvio padrão do modelo binomial, ou seja:

$$\boxed{\begin{aligned} \mu &= np \\ \sigma &= \sqrt{np(1-p)} \end{aligned}} \qquad \begin{aligned} (6.27) \\ (6.28) \end{aligned}$$

Exemplo 6.7 Historicamente, 10% dos pisos cerâmicos, que saem de uma linha de produção, têm algum defeito leve. Se a produção diária é de 1000 unidades, qual é a probabilidade de ocorrer mais de 120 itens defeituosos?

Pelas características do experimento, a variável aleatória Y = *número de defeituosos na amostra* tem distribuição binomial com parâmetros $n = 1000$ e $p = 0{,}1$. Verificamos, também, que as condições 6.25 e 6.26 estão satisfeitas, pois

a) $np = (1000)(0{,}1) = 100 \geq 5$ e

b) $n(1 - p) = (1000)(0{,}9) = 900 \geq 5$

Usando (6.27) e (6.28), temos:

$$\mu = np = 1000.(0{,}1) = 100 \text{ e}$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{1000 \cdot (0{,}1) \cdot (0{,}9)} = \sqrt{90}$$

Considere X uma variável aleatória normal com média $\mu = 100$ e variância $\sigma^2 = 90$. Então:

$$P(Y > 120) \approx P(X > 120)$$

Binomial com $n = 1.000$ e $p = 0{,}1$	Normal com $\mu = 100$ e $\sigma^2 = 90$

Valor padronizado: $z = \dfrac{x - \mu}{\sigma} = \dfrac{120 - 100}{\sqrt{90}} = 2{,}11$

Assim, $P(X > 120) = P(Z > 2{,}11) = 0{,}0174$

Correção de continuidade

Ao calcularmos probabilidades de eventos oriundos de experimentos binomiais como áreas sob uma curva normal, estamos fazendo uma aproximação de uma variável aleatória discreta, que só assume valores inteiros, para uma variável contínua, cujos eventos constituem intervalos de números reais. Nesse contexto, devemos fazer alguns ajustes, como ilustra o exemplo seguinte.

Exemplo 6.8 Seja Y o número de caras obtido em dez lançamentos de uma moeda honesta. Vamos calcular a probabilidade de obter quatro caras usando a distribuição normal.

Pelas características do experimento, Y tem distribuição binomial com $n = 10$ e $p = 0{,}5$. Então, a média e o desvio padrão são dados por:

$\mu =$

$$\sigma = \sqrt{np(1-p)} = \sqrt{10 \cdot (0{,}5) \cdot (1 - 0{,}5)} = \sqrt{2{,}5}$$

Considere o evento: *ocorrer quatro caras,* ou seja {y = 4}. Ao expressar este evento em termos de uma variável aleatória contínua X : N(5, *2,5),* devemos considerar um intervalo em torno do valor *4,* pois para variáveis aleatórias contínuas só faz sentido avaliar probabilidades em *intervalos.* O intervalo adequado, nesse caso, é construído pela subtração e soma de *meia unidade* ao valor quatro, ou seja, {3,5 < X < 4,5}, como ilustra a Figura 6.14.

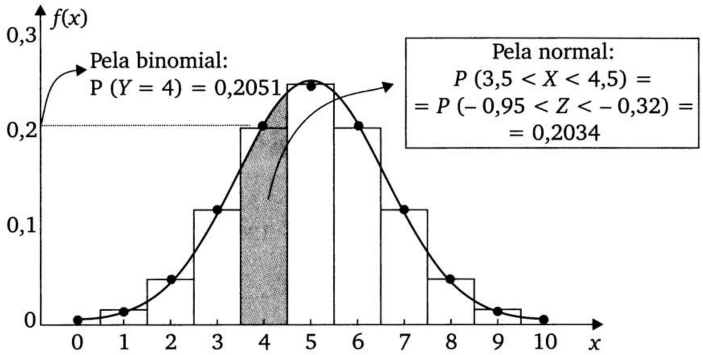


Figura 6.14 *Aproximação da probabilidade do evento {Y = 4} (em termos da distribuição binomial) para a probabilidade do evento {3,5* < X < 4,5} *(em termos da distribuição normal).*

O procedimento de subtrair e somar meia unidade para construir um intervalo em torno de valores inteiros é conhecido como *correção de continuidade.* Esta correção deve ser usada ao aproximar um evento de uma variável aleatória que só assume valores inteiros para um evento de uma variável aleatória contínua.

A Figura 6.15 ilustra as diversas situações possíveis de probabilidade associada a uma variável aleatória discreta, assumindo valores em {0, 1, 2, ...}, aproximada por probabilidade associada a uma variável aleatória contínua, a qual pode assumir qualquer valor real.

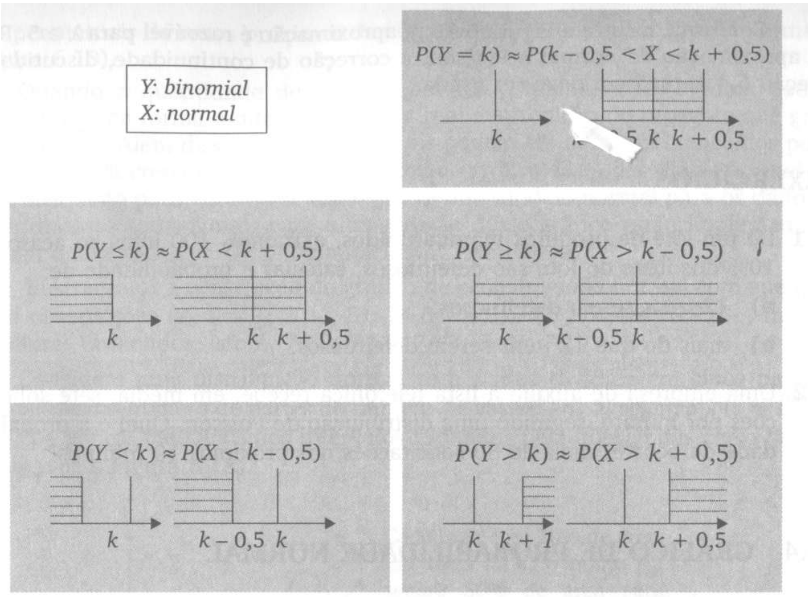


Figura 6.15 *Correção de continuidade ao aproximar uma variável aleatória discreta por uma variável aleatória contínua.*

6.3.2 Aproximação normal à Poisson

A distribuição de Poisson (Figura 6.16) também se aproxima da normal quando λ *é* grande. Como o valor esperado e a variância de uma Poisson são ambos iguais a λ, então, na aproximação normal, devemos usar:

$$\boxed{\begin{aligned}\mu &= \lambda \\ \sigma &= \sqrt{\lambda}\end{aligned}} \qquad \begin{aligned}(6.29)\\(6.30)\end{aligned}$$

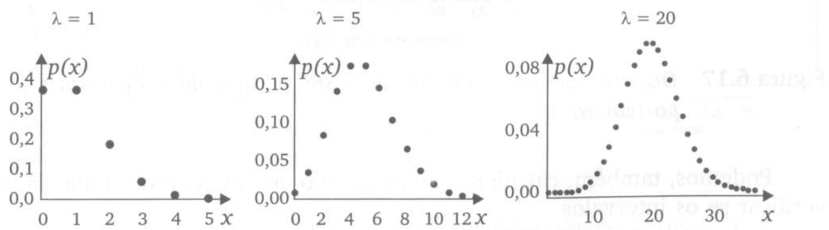


Figura 6.16 *Distribuições de Poisson para diferentes valores de λ.*

Conforme mostra a Figura 6.16, a aproximação é razoável para $\lambda \geq 5$. Para a aproximação da normal à Poisson, a correção de continuidade, discutida na Seção 6.4.1, também deve ser usada.

EXERCÍCIOS

11. De um lote de produtos manufaturados, extraímos 100 itens ao acaso. Se 10% dos itens do lote são defeituosos, calcular a probabilidade de:

a) 12 itens serem defeituosos;

b) mais do que 12 itens serem defeituosos.

12. Uma empresa de auxílio à lista telefônica recebe, em média, sete solicitações por minuto, segundo uma distribuição de Poisson. Qual é a probabilidade de ocorrer mais de 80 solicitações nos próximos 10 minutos?

6.4 GRÁFICO DE PROBABILIDADE NORMAL

Como veremos nos capítulos posteriores, muitos métodos estatísticos são desenvolvidos na suposição de que os dados provêm de uma distribuição normal. Quando o número de observações é grande, podemos construir um histograma e verificar se sua forma segue uma curva em forma de sino, sugerindo o modelo normal. É o caso da Figura 6.17.

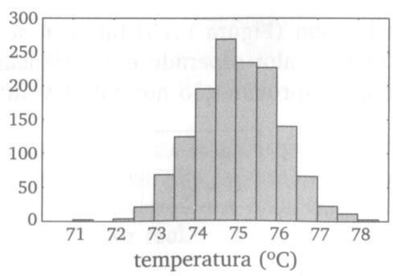


Figura 6.17 *Distribuição de freqüências de 1.389 leituras da temperatura de um pasteurizador.*

Podemos, também, calcular a média (x') e o desvio padrão (s) dos dados e verificar se os intervalos

$$\bar{x} \pm s, \ \bar{x} \pm 2s \text{ e } \bar{x} \pm 3s$$

têm percentuais de casos próximos dos esperados por uma distribuição normal (Figura 6.10).

Quando a quantidade de observações for pequena, o histograma pode apresentar uma forma muito diferente da real distribuição do processo que gerou os dados. Além disso, o cálculo de x' *e s* podem ser muito influenciados por algum valor discrepante. Em geral, o chamado *gráfico de probabilidade normal é* mais adequado para verificar a suposição de um modelo normal para os dados. Algoritmos computacionais para a construção desse gráfico estão implerpentados em quase todos pacotes computacionais estatísticos.[1]

Ilustraremos a construção do gráfico de probabilidade normal com apenas cinco observações (x_i i = 1, 2, ..., 5): 74,8; 74,0; 74,7; 74,4 e 75,9. Sejam $x_{(i)}$ os valores ordenados, isto é, 74,0; 74,4; 74,7; 74,8; 75,9.

Considere uma distribuição normal com a área dividida em cinco partes iguais (mesmo número de partes do número de valores, n). E sejam $q_{(i)}$ (i = 1, 2, ..., 5) os pontos medianos dos intervalos formados pela divisão das cinco áreas iguais (ver a Figura 6.18).

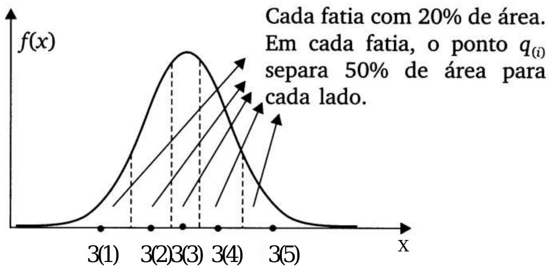
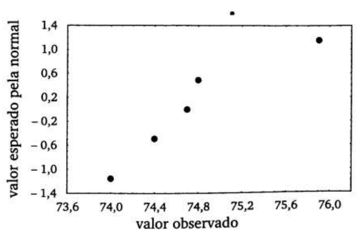


Figura 6.18 *Configuração de cinco pontos com as posições relativas mais verossímeis possíveis sob um modelo normal*

Se as cinco observações (74,0; 74,4; 74,7; 74,8; 75,9) provêm de uma distribuição normal, devemos esperar uma relação aproximadamente linear com os valores teóricos $q_{(i)}$. O gráfico de probabilidade normal compreende a apresentação dos pontos ($x_{(i)}$, $q_{(i)}$), num par de eixos cartesianos, conforme o gráfico ao lado. Nesse gráfico, os valores de $q_{(i)}$ foram padronizados (média zero e variância 1).



A Figura 6.19 apresenta dois gráficos de probabilidade normal. O gráfico da esquerda foi construído com 40 observações que aparentemente seguem uma distribuição normal. No gráfico da direita introduzimos um valor discrepante.

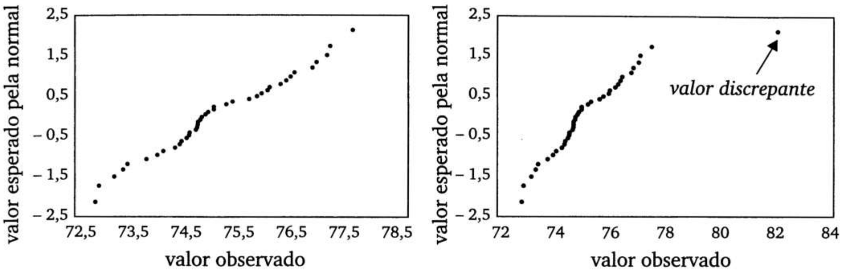


Figura 6.19 *Gráfico de probabilidade normal referente a 40 leituras de temperatura de um pasteurizador e o efeito de um valor discrepante.*

A Figura 6.20 mostra um gráfico de probabilidade normal construído com dados gerados por uma distribuição assimétrica, como mostrado do lado esquerdo da figura. Note que os pontos não estão aleatoriamente em torno de uma reta.

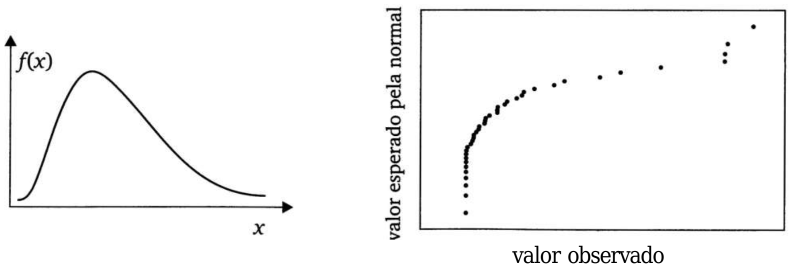


Figura 6.20 *Gráfico de probabilidade normal referente a 40 observações geradas por uma distribuição assimétrica.*

EXERCÍCIOS COMPLEMENTARES

13. O setor de manutenção de uma empresa fez um levantamento das falhas de um importante equipamento, constatando que há, em média, 0,75 falha por ano e que o tempo entre falhas segue uma distribuição exponencial. Qual é a probabilidade de o equipamento não falhar no próximo ano?

14. A vida útil de certo componente eletrônico é, em média, 10.000 horas e apresenta distribuição exponencial. Qual é a percentagem esperada de componentes que apresentarão falhas em menos de 10.000 horas?

15. A vida útil de certo componente eletrônico é, em media, 10.000 horas e apresenta distribuição exponencial. Após quantas horas se espera que 25% dos componentes tenham falhado?

16. Na manufatura de fios de linha para costura ocorre, em média, um Refeito a cada 100 metros de linha, segundo uma distribuição de Poisson.

a) Qual é a probabilidade de o próximo defeito ocorrer após 120 metros?

b) Quantos metros de linha poderão ser percorridos para que a probabilidade de aparecimento de algum defeito seja de 10%?

17. Num laticínio, a temperatura do pasteurizador deve ser de 75°C. Se a temperatura ficar inferior a 70°C, o leite poderá ficar com bactérias maléficas ao organismo humano. Observações do processo mostram que valores da temperatura seguem uma distribuição normal com média 75,4°C e desvio padrão 2,2°C.

a) Qual é a probabilidade da temperatura ficar inferior a 70°C?

b) Qual é a probabilidade de que, em 500 utilizações do pasteurizador, em mais do que cinco vezes a temperatura não atinja 70°C? Precisa supor distribuição normal.

18. O tempo para que um sistema computacional execute determinada tarefa é uma variável aleatória com distribuição normal, com *média 320* segundos e *desvio padrão de 7* segundos.

a) Qual é a probabilidade de a tarefa ser executada entre 310 e 330 segundos?

b) Se a tarefa é colocada para execução 200 vezes. Qual é a probabilidade de ela demorar mais do que 325 segundos em pelo menos 50 vezes?

19. **a)** Um exame de múltipla escolha consiste em dez questões, cada uma com quatro possibilidades de escolha. A aprovação exige, no mínimo, 50% de acertos. Qual é a probabilidade de aprovação se o candidato comparece ao exame sem saber absolutamente nada, apelando apenas para o "palpite"?

b) E se o exame tivesse 100 questões?

20. No horário de maior movimento, um sistema de banco de dados recebe, em média, 100 requisições por minuto, segundo uma distribuição de Poisson. Qual é a probabilidade de que no próximo minuto ocorram mais de 120 requisições? Use a aproximação normal com correção de continuidade.

21. Os dados históricos de uma rede de computadores sugerem que as conexões com essa rede, em horário normal, seguem uma distribuição de Poisson com média de cinco conexões por minuto. Calcule t_0, tal que se tenha probabilidade igual a 0,90 de que ocorra pelo menos uma conexão antes do tempo t_0.

22. O padrão de qualidade recomenda que os pontos impressos por uma impressora estejam entre 3,7 e 4,3 mm. Uma impressora imprime pontos, cujo diâmetro médio é igual a 4 mm e o desvio padrão é 0,19 mm. Suponha que o diâmetro dos pontos tenha distribuição normal.

a) Qual é a probabilidade do diâmetro de um ponto dessa impressora estar dentro do padrão?

b) Qual deveria ser o desvio padrão para que a probabilidade do item (a) atingisse 95%?

23. Certo tipo de cimento tem resistência à compressão com média de 5.800 kg/cm^2, e desvio padrão de 180 kg/cm^2, segundo uma distribuição normal. Dada uma amostra desse cimento, calcule as seguintes probabilidades:

a) resistência inferior a 5.600 kg/cm^2;

b) resistência entre 5.600 kg/cm^2 e 5.950 kg/cm^2;

c) resistência superior a 6.000 kg/cm^2, sabendo-se que ele já resistiu a 5.600 kg/cm^2.

d) se quer a garantia de que haja 95% de probabilidade de o cimento resistir a determinada pressão, qual deve ser o valor máximo dessa pressão?

24. Uma empresa fabrica dois tipos de monitores de vídeo. É suposto que as durabilidades deles seguem distribuições normais, sendo o monitor Ml com média de 6 anos e desvio padrão 2,3 anos; e o monitor M2 com média de 8 anos e desvio padrão 2,8 anos. Ml tem 2 anos de garantia e M2 tem 3 anos. A empresa lucra R$ 100,00 a cada Ml vendido e R$ 200,00 a cada M2 vendido, mas se deixarem de funcionar no período de garantia, a empresa perde R$ 300,00 (no caso de M1) e R$ 800,00 (no caso de M2). Em média, qual é o tipo de monitor que gera mais lucro?