Capítulo 7

Variáveis Aleatórias Contínuas

7.1 Introdução

Neste capítulo iremos estudar modelos probabilísticos para variáveis aleatórias contínuas, ou seja, variáveis para as quais os possíveis valores pertencem a um intervalo de números reais. A definição dada no capítulo anterior, para v.a. discreta, deve ser modificada como segue.

Definição. Uma função X, definida sobre o espaço amostral Ω e assumindo valores num intervalo de números reais, é dita uma *variável aleatória contínua*.

No Capítulo 2 vimos alguns exemplos de variáveis contínuas, como o salário de indivíduos, alturas etc. A característica principal de uma v.a. contínua é que, sendo resultado de uma mensuração, o seu valor pode ser pensado como pertencendo a um intervalo ao redor do valor efetivamente observado. Por exemplo, quando dizemos que a altura de uma pessoa é 175 cm, estamos medindo sua altura usando cm como unidade de medida e portanto o valor observado é, na realidade, um valor entre 174,5 cm e 175,5 cm.

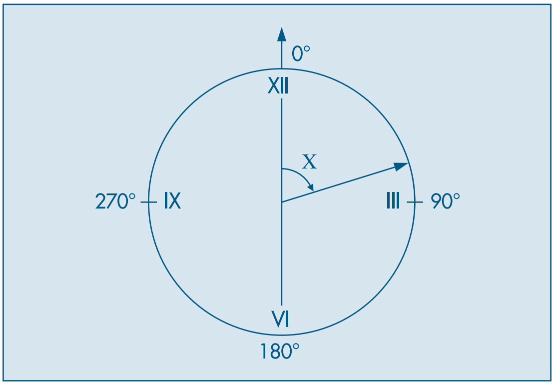
Vejamos um exemplo para motivar a discussão que se segue.

Exemplo 7.1. O ponteiro dos segundos de um relógio mecânico pode parar a qualquer instante, devido a algum defeito técnico, ou término da bateria, e vamos indicar por X o ângulo que esse ponteiro forma com o eixo imaginário passando pelo centro do mostrador e pelo número XII, conforme mostra a Figura 7.1.

Tabela 7.1: Distribuição uniforme discreta.

x	0°	6°	12°	18°	...	348°	354°
$p(x)$	1/60	1/60	1/60	1/60	...	1/60	1/60

Figura 7.1: Ilustração de uma v.a. X discreta.

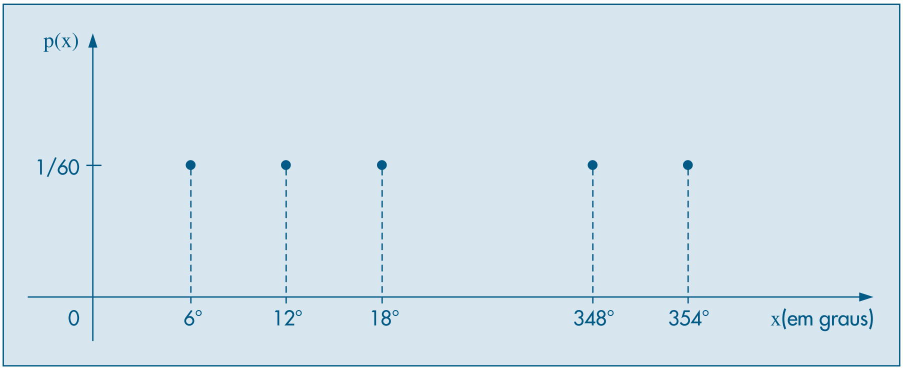


Medindo esse ângulo X em graus e lembrando que:

(i) o ponteiro deve dar 60 "saltos" (ele dá um salto em cada segundo) para completar uma volta;

(ii) acreditamos que o ponteiro tenha probabilidade igual de parar em qualquer ponto, então, a v.a. X tem distribuição uniforme discreta, com função de probabilidade dada pela Tabela 7.1 e representada graficamente na Figura 7.2.

Figura 7.2: Distribuição uniforme discreta.



Considerando esse mesmo problema com um relógio elétrico, para o qual o ponteiro dos segundos move-se *continuamente*, necessitamos de um outro modelo para representar a v.a. X. Primeiro, observamos que o conjunto dos possíveis valores de X não é mais um conjunto discreto de valores, pois X pode assumir qualquer valor do intervalo $[0,360) = = \{x \in \mathbb{R}: 0 \leqslant x < 360\}$. Em segundo lugar, como no caso do relógio mecânico, continuamos a acreditar que não exista uma região de preferência para o ponteiro parar. Como existem infinitos pontos nos quais o ponteiro pode parar, cada um com igual probabilidade, se fôssemos usar o mesmo método usado para a v.a. discreta uniforme, cada ponto teria probabilidade de ocorrer igual a zero. Assim não tem muito sentido falar na probabilidade de que o ângulo X seja igual a certo valor, pois essa probabilidade sempre será igual a zero. Entretanto, podemos determinar a probabilidade de que X esteja compreendido entre dois valores quaisquer. Por exemplo, usando a Figura 7.1 como referência, a probabilidade de o ponteiro parar no intervalo compreendido entre os números XII e III é 1/4, pois esse intervalo corresponde a 1/4 do intervalo total.

Podemos, pois, escrever

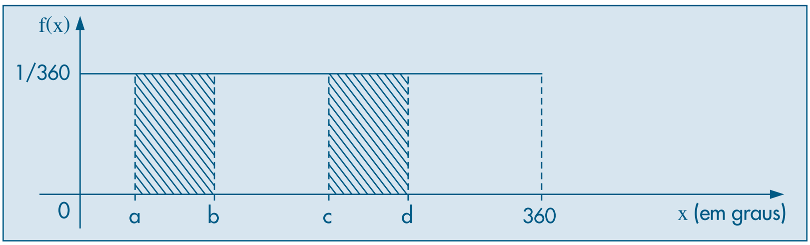
$$P(0^o \leqslant X \leqslant 90^o) = \frac{1}{4}.$$

Do mesmo modo, a probabilidade $P(120^o \leqslant X \leqslant 150^o) = 1/12$. Por menor que seja o intervalo, sempre poderemos calcular a probabilidade de o ponteiro parar num ponto qualquer desse intervalo. E é fácil verificar que, nesse caso, dados dois números a e b, tais que $0^o \leqslant a < b < 360^o$, a probabilidade de $X \in [a, b)$ é

$$P(a \leqslant X < b) = \frac{b-a}{360^o}$$

Através da divisão do intervalo $[0^o, 360^o)$ em pequenos subintervalos, podemos construir um histograma para as probabilidades da v.a. X (como fizemos para v.a contínuas no Capítulo 2). Ou ainda, como naquele capítulo, fazendo esses intervalos tenderem a zero, podemos construir o histograma alisado da v.a. X, apresentado na Figura 7.3.

Figura 7.3: Histograma alisado: distribuição uniforme contínua.



O histograma alisado da Figura 7.3 corresponde à seguinte função:

$$f(x) = \begin{cases} 0, & \text{se } x < 0^o \\ 1/360, & \text{se } 0^o \leqslant x < 360^o \\ 0, & \text{se } x \geqslant 360^o. \end{cases}$$

Como vimos na construção de histogramas, a área correspondente ao intervalo $[a, b)$ (hachurada na Figura 7.3) deve indicar a probabilidade de a variável estar entre a e b. Matematicamente, isso é expresso por meio da integral da função entre a e b; então,

$$P(a \leqslant X < b) = \int_a^b f(x)dx = \int_a^b \frac{1}{360}\, dx = \frac{b-a}{360},$$

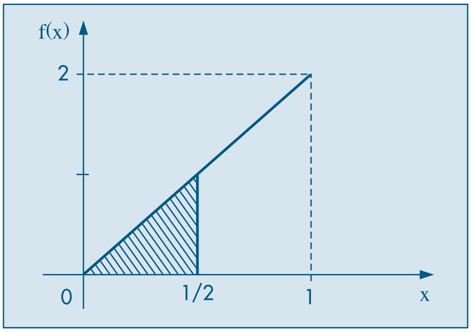
pois a integral definida de uma função entre dois pontos determina a área sob a curva representativa da função, compreendida entre esses dois pontos.

A função $f(x)$ é chamada *função densidade de probabilidade* (f.d.p.) da v.a. X.

Podemos construir modelos teóricos para variáveis aleatórias contínuas, escolhendo adequadamente as funções densidade de probabilidade. Teoricamente, qualquer função f, que seja não negativa e cuja área total sob a curva seja igual à unidade, caracterizará uma v.a. contínua.

Exemplo 7.2. Se $f(x) = 2x$, para $0 \leqslant x \leqslant 1$, e zero fora desse intervalo, vemos que $f(x) \geqslant 0$, para qualquer x, e a área sob o gráfico de $f(x)$ é unitária (verifique na Figura 7.4). Logo, a função f pode representar a função densidade de uma v.a. contínua X.

Figura 7.4: f.d.p. da v.a. X do Exemplo 7.2.



Para esse caso, $P(0 \leqslant X \leqslant 1/2)$ é igual à área do triângulo de base 1/2 e altura 1, hachurado na Figura 7.4; logo, a probabilidade em questão é

$$P(0 \leqslant X \leqslant 1/2) = \frac{1}{2}\left(\frac{1}{2} \times 1\right) = \frac{1}{4}.$$

Observamos, então, que a probabilidade de essa v.a. assumir um valor pertencente ao intervalo [0, 1/2) é menor que a probabilidade de a variável assumir um valor pertencente ao intervalo [1/2, 1).

A comparação das funções densidade dos dois últimos exemplos ajuda a entender seu significado. No primeiro exemplo, consideremos dois intervalos, $I_1 = [a, b)$ e $I_2 = [c, d)$, contidos no intervalo [0,360), com a mesma amplitude ($b - a = d - c$); então,

$$P(X \in I_1) = P(X \in I_2).$$

O mesmo não acontece no segundo exemplo: dados dois intervalos de mesma amplitude, aquele mais próximo de 1 irá apresentar maior probabilidade. Ou seja, a probabilidade de que a v.a. X assuma um valor num intervalo de amplitude fixa depende da posição do intervalo; existem regiões com maior *chance* de ocorrer, e o que determina esse fato é a função densidade de probabilidade. Portanto, a f.d.p. é um indicador da concentração de "massa" (probabilidade) nos possíveis valores de X. Convém ressaltar ainda que $f(x)$ não representa a probabilidade de ocorrência de algum evento. A área sob a curva entre dois pontos é que irá fornecer a probabilidade.

Problemas

1. Dada a função

$$f(x) = \begin{cases} 2e^{-2x}, \ x \geqslant 0 \\ \\ 0, \ x < 0, \end{cases}$$

   (a) Mostre que esta é uma f.d.p.
   (b) Calcule a probabilidade de $X > 10$.

2. Uma v.a. X tem distribuição triangular no intervalo $[0, 1]$ se sua f.d.p. for dada por

$$f(x) = \begin{cases} 0, & x < 0 \\ Cx, & 0 \leqslant x \leqslant 1/2 \\ C(1-x), & 1/2 \leqslant x \leqslant 1 \\ 0, & x > 1. \end{cases}$$

   (a) Qual valor deve ter a constante C?
   (b) Faça o gráfico de $f(x)$.
   (c) Determine $P(X \leqslant 1/2)$, $P(X > 1/2)$ e $P(1/4 \leqslant X \leqslant 3/4)$.

3. Suponha que estamos atirando dardos num alvo circular de raio 10 cm, e seja X a distância do ponto atingido pelo dardo ao centro do alvo. A f.d.p. de X é

$$f(x) = \begin{cases} kx, & \text{se } 0 \leqslant x \leqslant 10 \\ \\ 0, & \text{para os demais valores.} \end{cases}$$

   (a) Qual a probabilidade de acertar o centro do alvo, se esse for um círculo de 1 cm de raio?
   (b) Mostre que a probabilidade de acertar qualquer círculo concêntrico é proporcional à sua área.

4. Encontre o valor da constante c se

$$f(x) = \begin{cases} c/x^2, & x \geqslant 10 \\ 0, & x < 10 \end{cases}$$

   for uma densidade. Encontre $P(X > 15)$.

7.2 Valor Médio de uma Variável Aleatória Contínua

Do que foi visto até aqui, deduz-se que qualquer função $f(\cdot)$, não-negativa, tal que

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

define uma v.a. contínua X, ou seja, cria um modelo teórico para as freqüências relativas de uma v.a. contínua. A área compreendida entre dois valores, a e b, da abscissa x, sob a curva representativa de $f(x)$, dá a probabilidade (proporção teórica) da variável pertencer ao intervalo limitado pelos dois valores. Usando o conceito de integral, podemos escrever

$$P(a \leqslant X \leqslant b) = \int_a^b f(x)dx. \tag{7.1}$$

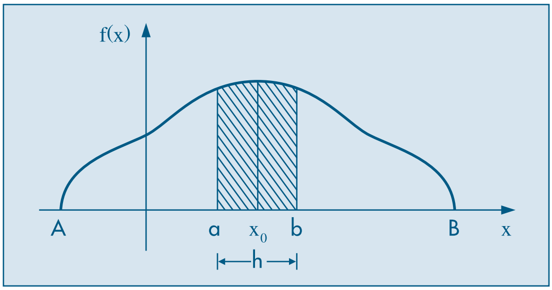
Vejamos agora como podemos definir a esperança (valor médio ou média) de uma v.a. contínua. Para isso, usaremos um artifício semelhante àquele usado na seção 3.1 para calcular a média das variáveis quantitativas, com os dados agrupados em classes. Lá substituímos todos os valores de um intervalo (classe) por um único valor aproximado (o ponto médio do intervalo), e agimos como se a variável fosse do tipo discreto. Aqui iremos repetir esse artifício.

Consideremos a v.a. X com função densidade $f(x)$ e dois pontos a e b, bem próximos, isto é, $h = b - a$ é pequeno, e consideremos x_0 o ponto médio do intervalo $[a, b]$. Observando a Figura 7.5 é fácil verificar que

$$P(a \leqslant X \leqslant b) \simeq h\, f(x_0), \tag{7.2}$$

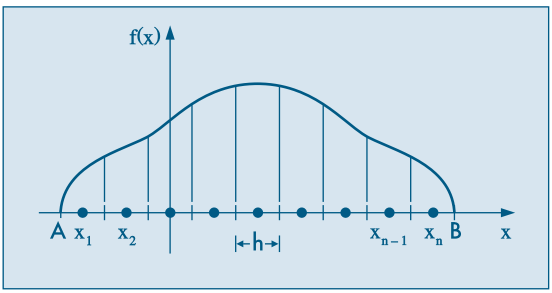
o que significa aproximar a área da parte hachurada pelo retângulo de base h e altura $f(x_0)$. É fácil ver que a aproximação melhora com h tendendo a zero.

Figura 7.5: Área hachurada representa $P(a \leqslant X \leqslant b)$.



Dividamos agora o intervalo $[A, B]$, onde $f(x) > 0$, em n partes de amplitudes iguais a $h = (B - A)/n$ (Figura 7.6) e consideremos os pontos médios desses intervalos, $x_1, x_2, ..., x_n$.

Figura 7.6: Partição do intervalo $[A, B]$.



Consideremos a v.a. Y_n, assumindo os valores $x_1, \ldots, x_n$ com as probabilidades

$$p_i = P(Y_n = x_i) \simeq f(x_i)h.$$

Dessa maneira, e de acordo com a definição de esperança, temos

$$E(Y_n) = \sum_{i=1}^{n} x_i p_i \approx \sum_{i=1}^{n} x_i\, f(x_i)h,$$

que será uma aproximação da esperança $E(X)$. Para determinar $E(X)$ com maior precisão, podemos aumentar o número de intervalos, diminuindo sua amplitude h. No limite, quando $h \to 0$, teremos o valor de $E(X)$. Definamos, pois,

$$E(X) = \lim_{n\to\infty} E(Y_n) = \lim_{n\to\infty} \sum_{i=1}^{n} x_i\, f(x_i)h. \tag{7.3}$$

Mas da definição de integral (veja Morettin *et al.*, 2005), temos que, se o limite (7.3) existe, ele define a integral de $x\,f(x)$ entre A e B, isto é,

$$E(X) = \int_A^B x f(x)dx. \tag{7.4}$$

Exemplo 7.3. Continuando com o Exemplo 7.2, observamos que, dividindo o intervalo [0, 1] em n subintervalos, teremos $h = 1/n$, $x_i = (2i - 1)/2n$ e $f(x_i) = (2i - 1)/n$, $i = 1, 2, \ldots, n$. Portanto,

$$E(Y_n) = \sum_{i=1}^{n} \left(\frac{2i-1}{2\,n}\right)\left(\frac{2i-1}{n}\right)\left(\frac{1}{n}\right) = \frac{1}{2n^3} \sum_{i=1}^{n} (2i-1)^2$$

$$= \frac{1}{2n^3} \left\{\frac{n(2n+1)(2n-1)}{3}\right\} = \frac{1}{6}\left(2 + \frac{1}{n}\right)\left(2 - \frac{1}{n}\right),$$

na qual usamos o conhecido resultado que dá a soma dos quadrados dos primeiros n números ímpares. Logo,

$$E(X) = \lim_{n\to\infty} \frac{1}{6}\left(2 + \frac{1}{n}\right)\left(2 - \frac{1}{n}\right) = \frac{2}{3}.$$

O mesmo resultado é obtido diretamente da relação (7.4):

$$E(X) = \int_0^1 (x)(2x)dx = \left[\frac{2x^3}{3}\right]_0^1 = \frac{2}{3}.$$

Exemplo 7.4. No caso do relógio elétrico do Exemplo 7.1, obtemos

$$E(X) = \int_0^{360} x\,\frac{1}{360}\,dx = \left[\frac{1}{360}\,\frac{x^2}{2}\right]_0^{360} = 180,$$

que é o valor esperado devido à distribuição uniforme das freqüências teóricas.

Como a função $f(x)$ é sempre não-negativa, podemos escrever a esperança como

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx. \tag{7.5}$$

A extensão do conceito de variância para v.a. contínuas é feita de maneira semelhante e o equivalente à expressão (6.2) é

$$\text{Var}(X) = E[(X - E(X))^2] = \int_{-\infty}^{\infty} (x - E(X))^2 f(x) dx. \tag{7.6}$$

Exemplo 7.5. Para os dois exemplos vistos anteriormente, teremos:

(i) Para o caso do relógio,

$$\text{Var}(X) = \int_0^{360} (x - 180)^2 \frac{1}{360} dx = \frac{1}{360}\left[\frac{x^3}{3} - \frac{360x^2}{2} + 180^2 x\right]_0^{360} = 10.800;$$

(ii) Para o Exemplo 7.2,

$$\text{Var}(X) = \int_0^1 \left(x - \frac{2}{3}\right)^2 2x dx = 2\left[\frac{x^4}{4} - \frac{4x^3}{9} + \frac{2x^2}{9}\right]_0^1 = \frac{1}{18}.$$

Como no caso de v.a. discretas, o desvio padrão de uma v.a. contínua X é definido como

$$DP(X) = \sqrt{\text{Var}(X)}, \tag{7.7}$$

que é dado na mesma unidade de medida do que X. Deixamos a cargo do leitor a verificação de que o seguinte resultado vale, como conseqüência de (7.6):

$$\text{Var}(X) = E(X^2) - [E(X)]^2. \tag{7.8}$$

Como frisamos no Capítulo 6, freqüentemente usaremos outros símbolos para indicar os parâmetros discutidos, a saber:

$$E(X) = \mu(X),$$
$$\text{Var}(X) = \sigma^2(X),$$
$$DP(X) = \sigma(X),$$

ou simplesmente μ, σ^2 e σ, respectivamente, se não houver possibilidade de confusão.

7.3 Função de Distribuição Acumulada

Dada uma v.a. X com função densidade de probabilidade $f(x)$, podemos definir a sua função de distribuição acumulada, $F(x)$, do mesmo modo como foi definida no Capítulo 6:

$$F(x) = P(X \leqslant x), -\infty < x < \infty. \tag{7.9}$$

De (7.1) segue-se que

$$F(x) = \int_{-\infty}^{x} f(t) dt, \tag{7.10}$$

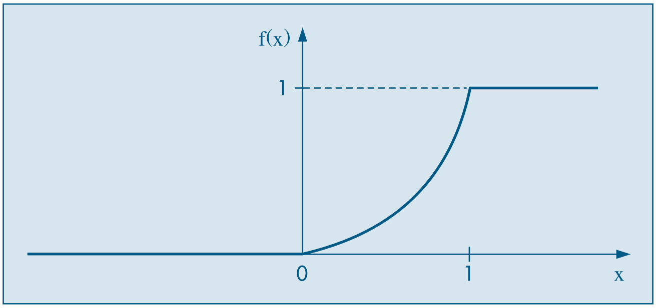
para todo real x.

Exemplo 7.6. Retomemos o Exemplo 7.2. Temos

$$F(x) = \begin{cases} 0, & \text{se } x < 0 \\ \int_0^x 2tdt = x^2, & \text{se } 0 \leqslant x < 1 \\ \int_0^1 2tdt + \int_1^x 0dt = 1, & \text{se } x \geqslant 1. \end{cases}$$

O gráfico de $F(x)$ está na Figura 7.7.

Figura 7.7: f.d.a. da v.a. X do Exemplo 7.6.



De (7.9), vemos que $0 \leqslant F(x) \leqslant 1$, para todo x real; além disso, $F(x)$ é não-decrescente e possui as duas seguintes propriedades:

(i) $\lim_{x \to -\infty} F(x) = 0$,

(ii) $\lim_{x \to \infty} F(x) = 1$.

No Exemplo 7.6 temos, efetivamente, $F(x) = 0$, para $x < 0$ e $F(x) = 1$, para $x \geqslant 1$.

Para v.a. contínuas, o seguinte resultado é importante.

Proposição 7.1. Para todos os valores de x para os quais $F(x)$ é derivável temos

$$F'(x) = \frac{dF(x)}{dx} = f(x).$$

Vamos usar esse resultado no exemplo a seguir.

Exemplo 7.7. Suponha que

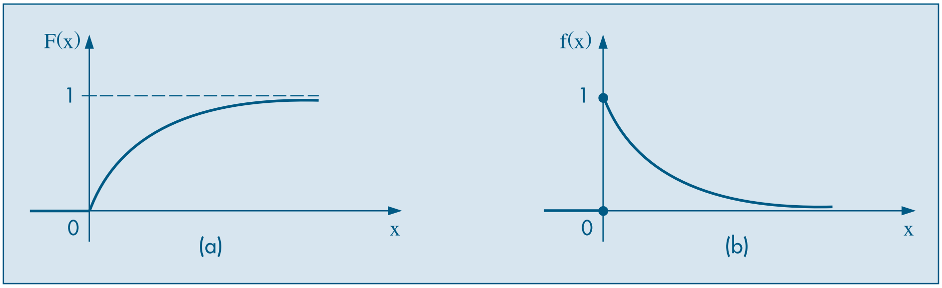
$$F(x) = \begin{cases} 0, & \text{se } x < 0 \\ 1 - e^{-x}, & \text{se } x \geqslant 0 \end{cases}$$

seja a f.d.a. de uma v.a. X. Então,

$$f(x) = \begin{cases} 0, & \text{se } x < 0 \\ e^{-x}, & \text{se } x \geqslant 0. \end{cases}$$

Na Figura 7.8 temos os gráficos dessas duas funções. Veremos que f(x) é um caso especial da densidade exponencial, a ser estudada na seção 7.4.3.

Figura 7.8: Distribuição exponencial ($\beta = 1$) (a) f.d.a. (b) f.d.p.



Se a e b forem dois números reais quaisquer,

$$P(a < X \leqslant b) = F(b) - F(a). \tag{7.11}$$

Esse resultado não será afetado se incluirmos ou não os extremos a e b na desigualdade entre parênteses.

Problemas

5. Calcule a esperança, a variância e a f.d.a. da v.a. X do Problema 2.

6. Determine a esperança e a variância da v.a. cuja f.d.p. é

$$f(x) = \begin{cases} \text{sen}\, x, & 0 \leqslant x \leqslant \pi/2 \\ 0, & \text{caso contrário.} \end{cases}$$

7. Calcule a média da v.a. X do Problema 4.

8. A v.a. contínua X tem f.d.p.

$$f(x) = \begin{cases} 3x^2, & -1 \leqslant x \leqslant 0 \\ 0, & \text{caso contrário.} \end{cases}$$

(a) Se b for um número que satisfaz $-1 < b < 0$, calcule $P(X > b \mid X < b/2)$.
(b) Calcule $E(X)$ e $\text{Var}(X)$.

9. Certa liga é formada pela mistura fundida de dois metais. A liga resultante contém certa porcentagem de chumbo, X, que pode ser considerada uma v.a. com f.d.p.

$$f(x) = \frac{3}{5}\, 10^{-5}\, x(100 - x),\ 0 \leqslant x \leqslant 100.$$

Suponha que L, o lucro líquido obtido na venda dessa liga (por unidade de peso), seja dado por $L = C_1 + C_2X$. Calcule $E(L)$, o lucro esperado por unidade.

10. A demanda diária de arroz num supermercado, em centenas de quilos, é uma v.a. com f.d.p.

$$f(x) = \begin{cases} 2x/3, & \text{se } 0 \leqslant x < 1 \\ -x/3 + 1, & \text{se } 1 \leqslant x < 3 \\ 0, & \text{se } x < 0 \text{ ou } x > 3. \end{cases}$$

(a) Qual a probabilidade de se vender mais do que 150 kg, num dia escolhido ao acaso?
(b) Em 30 dias, quanto o gerente do supermercado espera vender?
(c) Qual a quantidade de arroz que deve ser deixada à disposição dos clientes diariamente para que não falte arroz em 95% dos dias?

11. Suponha que X tenha f.d.p. $f(x)$ do Problema 1. Calcule $E(X)$ e $\text{Var}(X)$.

12. Seja X com densidade

$$f(x) = \begin{cases} c(1-x^2), & \text{se } -1 \leqslant x \leqslant 1 \\ 0, & \text{caso contrário.} \end{cases}$$

Calcule a média e a variância de X.

7.4 Alguns Modelos Probabilísticos para Variáveis Aleatórias Contínuas

De modo geral, podemos dizer que as v.a. cujos valores resultam de algum processo de mensuração são v.a. contínuas. Alguns exemplos são:

(a) o peso ou a altura das pessoas de uma cidade;
(b) a demanda diária de arroz num supermercado;
(c) o tempo de vida de uma lâmpada;
(d) o diâmetro de rolamentos de esferas; e
(e) erros de medidas em geral, resultantes de experimentos em laboratórios.

Dada uma v.a. contínua X, interessa saber qual a f.d.p. de X. Alguns modelos são freqüentemente usados para representar a f.d.p. de v.a. contínuas. Alguns dos mais utilizados serão descritos a seguir e, para uniformizar o estudo desses modelos, iremos em cada caso analisar:

(a) definição;
(b) gráfico da f.d.p.;
(c) momentos: $E(X)$, $\text{Var}(X)$;
(d) função de distribuição acumulada (f.d.a.).

Outros modelos serão apresentados na seção 7.7.

7.4.1 O Modelo Uniforme

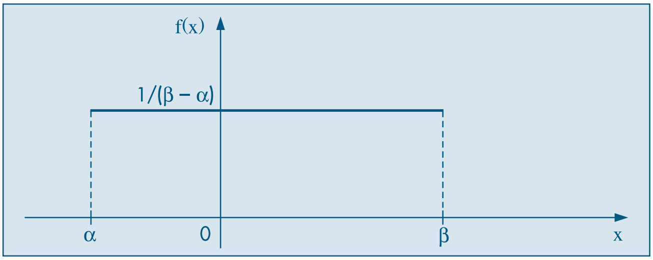
O modelo uniforme é uma generalização do modelo estudado no Exemplo 7.1 e é o modelo mais simples para v.a. contínuas.

(a) **Definição.** A v.a. X tem distribuição uniforme no intervalo $[\alpha, \beta]$ se sua f.d.p. é dada por

$$f(x;\alpha,\beta)=\begin{cases}\dfrac{1}{\beta-\alpha}, & \text{se } \alpha\leqslant x\leqslant\beta,\\ 0, & \text{caso contrário.}\end{cases} \tag{7.12}$$

(b) **Gráfico.** A Figura 7.9 representa a função dada por (7.12).

Figura 7.9: Distribuição uniforme no intervalo $[\alpha, \beta]$.



(c) **Momentos.** Pode-se mostrar (veja o Problema 29) que

$$E(X)=\frac{\alpha+\beta}{2}, \tag{7.13}$$

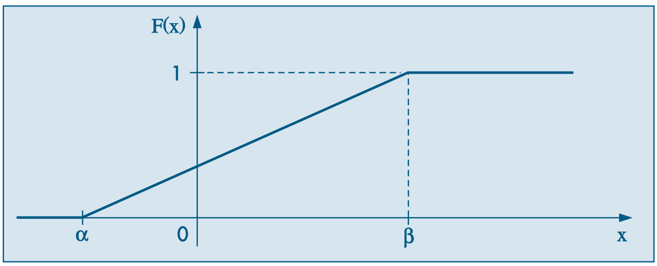
$$\mathrm{Var}(X)=\frac{(\beta-\alpha)^2}{12}. \tag{7.14}$$

(d) **F.d.a.** A função de distribuição acumulada da uniforme é fácil de ser encontrada (veja o Problema 29):

$$F(x)=P(X\leqslant x)=\int_{-\infty}^{x}f(x)dx=\begin{cases}0, & \text{se } x<\alpha\\ \dfrac{x-\alpha}{\beta-\alpha}, & \text{se } \alpha\leqslant x<\beta\\ 1, & \text{se } x\geqslant\beta,\end{cases} \tag{7.15}$$

cujo gráfico está na Figura 7.10.

Figura 7.10: f.d.a. de uma v.a. uniforme no intervalo $[\alpha, \beta]$.



Assim, para dois valores quaisquer c e d, $c < d$, teremos

$$P(c < X \leqslant d) = F(d) - F(c),$$

que é obtida facilmente de (7.15).

Usaremos a notação

$$X \sim \mathcal{U}(\alpha, \beta)$$

para indicar que a v.a. X tem distribuição uniforme no intervalo $[\alpha, \beta]$.

Exemplo 7.8. Um caso particular bastante interessante é aquele em que $\alpha = -1/2$ e $\beta = 1/2$. lndicando essa v.a. por U, teremos

$$f(u) = \begin{cases} 1, & \text{se } -1/2 \leqslant u \leqslant 1/2 \\ 0, & \text{caso contrário.} \end{cases}$$

Nessa situação temos que

$$E(U) = 0, \ \mathrm{Var}(U) = 1/12$$

e a f.d.a. é dada por

$$F_U(u) = \begin{cases} 0, & \text{se } u < -1/2 \\ u + 1/2, & \text{se } -1/2 \leqslant u < 1/2 \\ 1, & \text{se } u > 1/2. \end{cases}$$

Por exemplo,

$$P(-1/4 \leqslant \mathrm{U} \leqslant 1/4) = F_U(1/4) - F_U(-1/4) = 1/2.$$

Se quiséssemos facilitar o nosso trabalho, poderíamos tabelar os valores da f.d.a para essa variável U. Devido à simetria da área em relação a $x = 0$, poderíamos construir uma tabela indicando a função $G(u)$, tal que

$$G(u) = P(0 \leqslant U \leqslant u)$$

para alguns valores de u (veja o Problema 30).

Dada uma v.a. uniforme X qualquer, com parâmetros α e β, podemos definir a v.a. U como

$$U = \frac{X - \dfrac{\beta + \alpha}{2}}{\beta - \alpha}. \tag{7.16}$$

Segue-se que a transformação (7.16) leva uma uniforme no intervalo $[\alpha, \beta]$ numa uniforme no intervalo $[-1/2, 1/2]$ e para dois números quaisquer c e d, com $c < d$,

$$P(c < X \leqslant d) = F(d) - F(c) = P\left(\frac{c - \dfrac{\beta+\alpha}{2}}{\beta - \alpha} < U \leqslant \frac{d - \dfrac{\beta+\alpha}{2}}{\beta - \alpha}\right) = F_U\left(\frac{d - \dfrac{\beta+\alpha}{2}}{\beta - \alpha}\right) - F_U\left(\frac{c - \dfrac{\beta+\alpha}{2}}{\beta - \alpha}\right).$$

Artifícios semelhantes a esse são muito úteis na construção de tabelas e programas para cálculos de probabilidades referentes a famílias de modelos.

Um outro caso importante é para $\alpha = 0$ e $\beta = 1$. Um número aleatório é um valor gerado de uma v.a. com distribuição uniforme no intervalo [0, 1]. Veja Capítulo 9.

7.4.2 O Modelo Normal

Vamos introduzir, agora, um modelo fundamental em probabilidades e inferência estatística. Suas origens remontam a Gauss em seus trabalhos sobre erros de observações astronômicas, por volta de 1810, donde o nome de distribuição *gaussiana* para tal modelo.

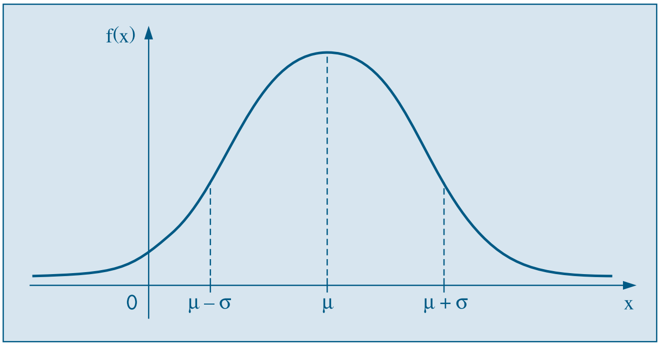
(a) **Definição.** Dizemos que a v.a. X tem *distribuição normal* com parâmetros μ e σ^2, $-\infty < \mu < +\infty$ e $0 < \sigma^2 < \infty$, se sua densidade é dada por

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty. \tag{7.17}$$

Claramente, $f(x; \mu, \sigma^2) \geqslant 0$, para todo x e pode-se provar que $\int_{-\infty}^{\infty} f(x; \mu, \sigma^2)\, dx = 1$. Veja o Problema 60.

(b) **Gráfico.** A Figura 7.11 ilustra uma particular *curva normal*, determinada por valores particulares de μ e σ^2.

Figura 7.11: f.d.p. de uma v.a. normal com média μ e desvio padrão σ.



(c) **Momentos.** Pode-se demonstrar que (veja o Problema 32):

$$E(X) = \mu, \tag{7.18}$$

$$\mathrm{Var}(X) = \sigma^2. \tag{7.19}$$

Além disso, $f(x; \mu; \sigma^2) \to 0$, quando $x \to \pm\infty$, $\mu - \sigma$ e $\mu + \sigma$ são pontos de inflexão de $f(x; \mu, \sigma^2)$, $x = \mu$ é ponto de máximo de $f(x; \mu, \sigma^2)$, e o valor máximo é $1/\sigma\sqrt{2\pi}$. A densidade $f(x; \mu, \sigma^2)$ é simétrica em relação à reta $x = \mu$, isto é,

$$f(\mu + x; \mu, \sigma^2) = f(\mu - x; \mu, \sigma^2), \tag{7.20}$$

para todo x real.

Para simplificar a notação, denotaremos a densidade da normal simplesmente por $f(x)$ e escreveremos, simbolicamente,

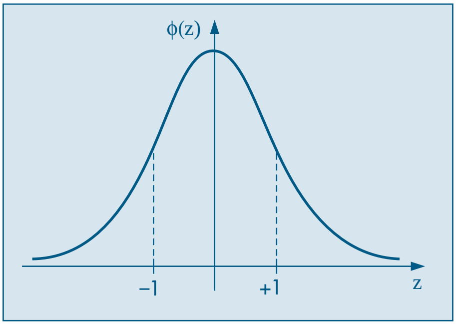
$$X \sim N(\mu, \sigma^2).$$

Quando $\mu = 0$ e $\sigma^2 = 1$, temos uma distribuição *padrão* ou *reduzida*, ou brevemente $N(0,1)$. Para essa a função densidade reduz-se a

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad -\infty < z < \infty. \tag{7.21}$$

O gráfico da normal padrão está na Figura 7.12.

Figura 7.12: f.d.p. de uma v.a. normal padrão: $Z \sim N(0, 1)$.



Se $X \sim N(\mu; \sigma^2)$, então a v.a. definida por

$$Z = \frac{X - \mu}{\sigma}, \tag{7.22}$$

terá média zero e variância 1 (prove esses fatos). O que não é tão fácil mostrar é que Z também tem distribuição normal. Isso não será feito aqui.

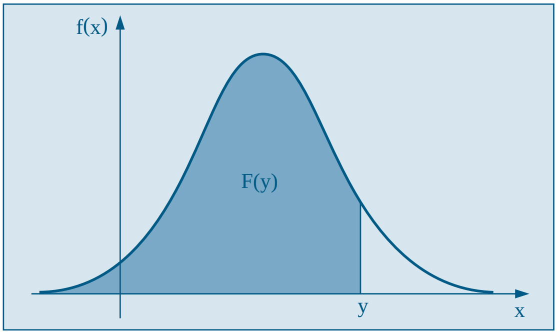
A transformação (7.22) é fundamental para calcularmos probabilidades relativas a uma distribuição normal qualquer.

(d) **F.d.a.** A f.d.a. $F(y)$ de uma v.a. normal X, com média μ e variância σ^2 é obtida integrando-se (7.17) de $-\infty$ até y, ou seja,

$$F(y) = \int_{-\infty}^{y} f(x; \mu, \sigma^2)dx, \quad y \in I\!R. \tag{7.23}$$

A integral (7.23) corresponde à área, sob $f(x)$, desde $-\infty$ até y, como ilustra a Figura 7.13.

Figura 7.13: Representação gráfica de F(y) como área.

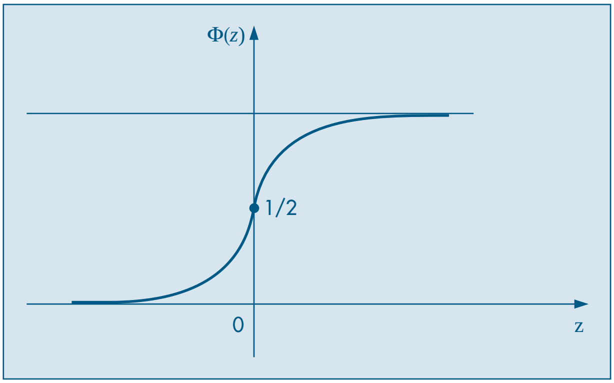


No caso específico da normal padrão, utilizamos a seguinte notação, que é universal:

$$\Phi(y) = \int_{-\infty}^{y} \phi(z)dz = 1/\sqrt{2\pi} \int_{-\infty}^{y} e^{-z^2/2}\, dz. \tag{7.24}$$

O gráfico de $\Phi(z)$ é ilustrado na Figura 7.14.

Figura 7.14: f.d.a. da normal padrão.

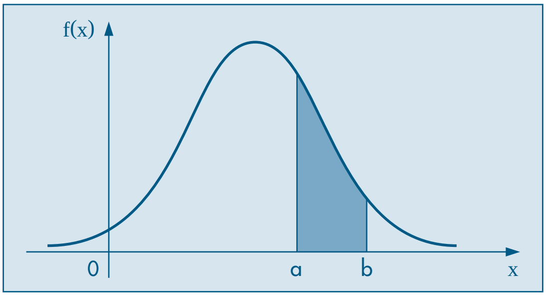


Suponha, então, que $X \sim N(\mu, \sigma^2)$ e que queiramos calcular

$$P(a < X < b) = \int_{a}^{b} f(x)dx, \tag{7.25}$$

onde $f(x)$ é dada por (7.17). Ver Figura 7.15.

Figura 7.15: Ilustração gráfica da $P(a \leqslant X \leqslant b)$ para uma v.a. normal.



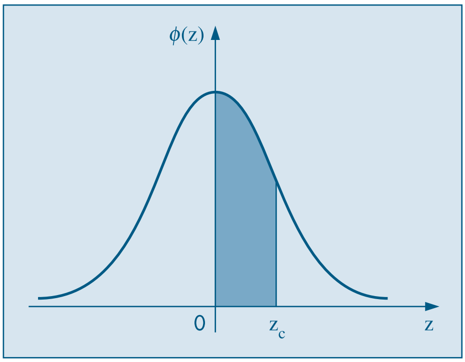
A integral (7.25) não pode ser calculada analiticamente, e portanto a probabilidade indicada só poderá ser obtida, aproximadamente, por meio de integração numérica. No entanto, para *cada* valor de μ e *cada* valor de σ, teríamos de obter $P(a < X < b)$ para diversos valores de a e b. Essa tarefa é facilitada através do uso de (7.22), de sorte que somente é necessário construir uma tabela para a distribuição normal padrão.

Vejamos, então, como obter probabilidades a partir da Tabela III. Essa tabela dá as probabilidades sob uma curva normal padrão, que nada mais são do que as correspondentes áreas sob a curva. A Figura 7.16 ilustra a probabilidade fornecida pela tabela, a saber,

$$P(0 \leqslant Z \leqslant z_c),$$

onde $Z \sim N(0,1)$.

Figura 7.16: $P(0 \leqslant Z \leqslant z_c)$ fornecido pela Tabela III.



Se tomarmos, por exemplo, $z_c = 1{,}73$, segue-se que

$$P(0 \leqslant Z \leqslant 1{,}73) = 0{,}4582.$$

Calculemos mais algumas probabilidades (Figura 7.17):

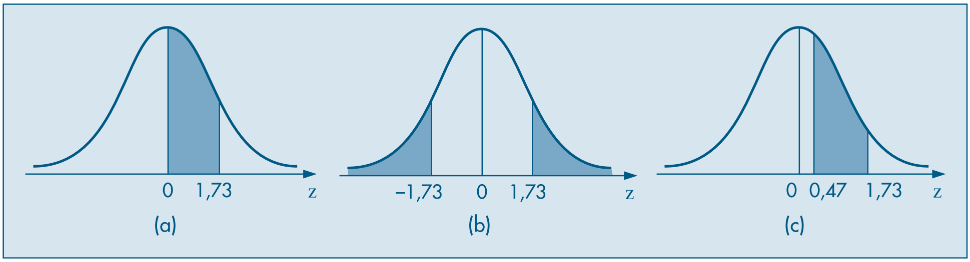
(a) $P(-1{,}73 \leqslant Z \leqslant 0) = P(0 \leqslant Z \leqslant 1{,}73) = 0{,}4582$, devido à simetria da curva.

(b) $P(Z \geqslant 1{,}73) = 0{,}5 - P(0 \leqslant Z \leqslant 1{,}73) = 0{,}5 - 0{,}4582 = 0{,}0418$, pois $P(Z \geqslant 0) = 0{,}5 = P(Z \leqslant 0)$.

(c) $P(Z < -1{,}73) = P(Z > 1{,}73) = 0{,}0418$.

(d) $P(0{,}47 \leqslant Z \leqslant 1{,}73) = P(0 \leqslant Z \leqslant 1{,}73) - P(0 \leqslant Z \leqslant 0{,}47) =$
$= 0{,}4582 - 0{,}1808 = 0{,}2774$.

Figura 7.17: Ilustração do cálculo de probabilidades para a $N(0,1)$.



Suponha, agora, que X seja uma v.a. $N(\mu, \sigma^2)$, com $\mu = 3$ e $\sigma^2 = 16$, e queiramos calcular $P(2 \leqslant X \leqslant 5)$. Utilizando (7.22), temos

$$P(2 \leqslant X \leqslant 5) = P\left(\frac{2-\mu}{\sigma} \leqslant \frac{X-\mu}{\sigma} \leqslant \frac{5-\mu}{\sigma}\right)$$

$$= P\left(\frac{2-3}{4} \leqslant Z \leqslant \frac{5-3}{4}\right) = P\left(-\frac{1}{4} \leqslant Z \leqslant \frac{1}{2}\right).$$

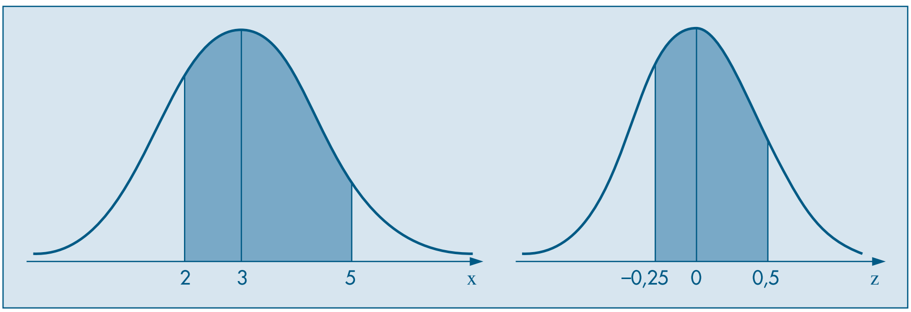
Portanto, a probabilidade de que X esteja entre 2 e 5 é igual à probabilidade de que Z esteja entre $-0{,}25$ e $0{,}5$ (Figura 7.18). Utilizando a Tabela III, vemos que

$$P(-0{,}25 \leqslant Z \leqslant 0{,}5) = 0{,}0987 + 0{,}1915 = 0{,}2902,$$

ou seja,

$$P(2 \leqslant X \leqslant 5) = 0{,}2902.$$

Figura 7.18: Ilustração do cálculo de $P(2 \leqslant X \leqslant 5)$ para a v.a. $N(3, 16)$.



Exemplo 7.9. Os depósitos efetuados no Banco da Ribeira durante o mês de janeiro são distribuídos normalmente, com média de \$10.000,00 e desvio padrão de \$1.500,00. Um depósito é selecionado ao acaso dentre todos os referentes ao mês em questão. Encontrar a probabilidade de que o depósito seja:

(a) \$10.000,00 ou menos;

(b) pelo menos \$10.000,00;

(c) um valor entre \$12.000,00 e \$15.000,00;

(d) maior do que \$20.000,00.

Temos que $\mu = 10.000$ e $\sigma = 1.500$. Seja a v.a. X = depósito.

(a) $P(X \leqslant 10.000) = P\left(Z \leqslant \dfrac{10.000 - 10.000}{1.500}\right) = P(Z \leqslant 0) = 0{,}5.$

(b) $P(X \geqslant 10.000) = P(Z \geqslant 0) = 0{,}5.$

(c) $P(12.000 < X < 15.000) = P\left(\dfrac{12.000 - 10.000}{1.500} < Z < \dfrac{15.000 - 10.000}{1.500}\right)$

$= P(4/3 < Z < 10/3) = P(1{,}33 < Z < 3{,}33) = 0{,}09133.$

(d) $P(X > 20.000) = P\left(Z > \dfrac{20.000 - 10.000}{1.500}\right) = P(Z > 6{,}67) \simeq 0.$

7.4.3 O Modelo Exponencial

Outra distribuição importante e que tem aplicações em confiabilidade de sistemas, assunto de que já tratamos brevemente no Capítulo 5, é a exponencial.

(a) **Definição.** A v.a. T tem *distribuição exponencial* com parâmetro $\beta > 0$ se sua f.d.p. tem a forma

$$f(t;\beta) = \begin{cases} \dfrac{1}{\beta}\, e^{-t/\beta}, & \text{se } t \geqslant 0 \\ 0, & \text{se } t < 0. \end{cases} \tag{7.26}$$

Escreveremos, brevemente,

$$T \sim \text{Exp}(\beta).$$

(b) **Gráfico.** O gráfico de $f(t;\beta) = f(t)$ está ilustrado na Figura 7.8 (b), com $\beta = 1$.

(c) **Momentos.** Usando integração por partes, pode-se demonstrar que (veja o Problema 41):

$$E(T) = \beta, \tag{7.27}$$

$$\text{Var}(T) = \beta^2. \tag{7.28}$$

Exemplo 7.10. O tempo de vida (em horas) de um transistor pode ser considerado uma v.a com distribuição exponencial com $\beta = 500$. Segue-se que a vida média do transistor é $E(T) = 500$ horas e a probabilidade de que ele dure mais do que a média é

$$P(T > 500) = \int_{500}^{\infty} f(t)dt = 1/500 \int_{500}^{\infty} e^{-t/500}\, dt$$

$$= 1/500\, [-500e^{-t/500}]_{500}^{\infty} = e^{-1} = 0{,}3678.$$

(d) **F.d.a.** Usando a definição (7.10), obtemos

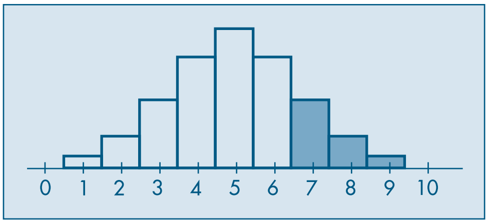
$$F(t) = \begin{cases} 0, & \text{se } t < 0 \\ 1 - e^{-t/\beta}, & \text{se } t \geqslant 0. \end{cases} \tag{7.29}$$

O gráfico de $F(t)$ está na Figura 7.8 (a), com $\beta = 1$.

7.5 Aproximação Normal à Binomial

Suponha que a v.a. Y tenha uma distribuição binomial com parâmetros $n = 10$ e $p = 1/2$ e queiramos calcular $P(Y \geqslant 7)$. Embora seja uma v.a. discreta, vimos no Capítulo 2 que é possível representá-la por meio de um histograma, como na Figura 7.19. Vemos que $P(Y = 7)$ é igual à área do retângulo de base unitária e altura igual a $P(Y = 7)$, similarmente para $P(Y = 8)$ etc. Logo, $P(Y \geqslant 7)$ é igual à soma das áreas dos retângulos hachurados na Figura 7.19.

Figura 7.19: $(P(Y \geqslant 7)$ para $Y \sim b(10, 1/2)$.



A idéia é aproximar tal área pela área sob uma curva normal, à *direita* de 6,5. Qual curva normal? Parece razoável considerar aquela normal de média

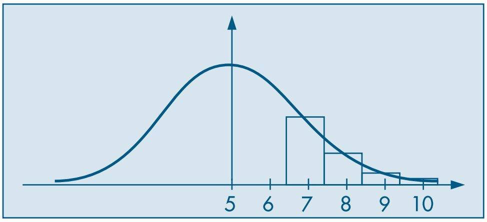
$$\mu = np = 10 \times \frac{1}{2} = 5$$

e variância

$$\sigma^2 = np(1 - p) = 10 \times \frac{1}{2} \times \frac{1}{2} = 2{,}5.$$

Veja a Figura 7.20.

Figura 7.20: Aproximação de $P(Y \geqslant 7)$ pela área sob a $N(5; 2{,}5)$.



Chamando X tal variável, com distribuição normal,

$$P(Y \geqslant 7) \simeq P(X \geqslant 6{,}5) = P\left(\frac{X-\mu}{\sigma} \geqslant \frac{6{,}5-\mu}{\sigma}\right)$$

$$P\left(Z \geqslant \frac{6{,}5-5}{\sqrt{2{,}5}}\right) = P(Z \geqslant 0{,}94) = 0{,}174,$$

onde Z é, como sempre, $N(0,\ 1)$. Utilizando a Tabela I, vemos que a probabilidade verdadeira é 0,172.

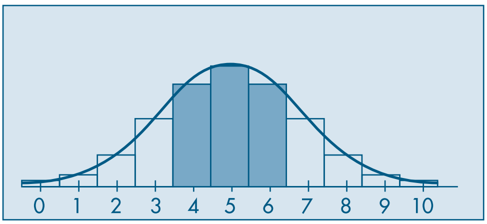
Vamos calcular agora $P(3 < Y \leqslant 6) = P(Y = 4) + P(Y = 5) + P(Y = 6)$. Vemos, através da Figura 7.21, que a aproximação a ser feita deve ser

$$P(3 < Y \leqslant 6) \simeq P(3{,}5 \leqslant X \leqslant 6{,}5) = P\left(\frac{3{,}5-5}{1{,}58} \leqslant Z \leqslant \frac{6{,}5-6}{1{,}58}\right)$$

$$= P(-0{,}94 \leqslant Z \leqslant 0{,}94) = 0{,}653,$$

ao passo que a probabilidade verdadeira é 0,656.

Figura 7.21: Aproximação de $P(3 < Y \leqslant 6)$.



A justificativa formal de tal aproximação é dada pelo chamado Teorema Limite Central, que será visto no Capítulo 10. A aproximação é boa quando $np > 5$ e $n(1-p) > 5$.

Problemas

13. A temperatura T de destilação do petróleo é crucial na determinação da qualidade final do produto. Suponha que T seja considerada uma v.a. com distribuição uniforme no intervalo (150, 300). Suponha que o custo para produzir um galão de petróleo

seja C_1 reais. Se o óleo for destilado a uma temperatura inferior a 200°, o produto obtido é vendido a C_2 reais; se a temperatura for superior a 200°, o produto é vendido a C_3 reais.

(a) Fazer o gráfico da f.d.p. de T.

(b) Qual o lucro médio por galão?

14. Se $X \sim N(10, 4)$, calcular:

(a) $P(8 < X < 10)$,
(c) $P(X > 10)$,
(b) $P(9 \leqslant X \leqslant 12)$,
(d) $P(X < 8 \text{ ou } X > 11)$.

15. Para $X \sim N(100, 100)$, calcule:

(a) $P(X < 115)$,

(b) $P(X \geqslant 80)$,

(c) $P(|X - 100| \leqslant 10)$,

(d) o valor a, tal que $P(100 - a \leqslant X \leqslant 100 + a) = 0{,}95$.

16. Para a v.a. $X \sim N(\mu, \sigma^2)$, encontre:

(a) $P(X \leqslant \mu + 2\sigma)$,

(b) $P(|X - \mu| \leqslant \sigma)$,

(c) o número a tal que $P(\mu - a\sigma \leqslant X \leqslant \mu + a\sigma) = 0{,}99$,

(d) o número b tal que $P(X > b) = 0{,}90$.

17. As alturas de 10.000 alunos de um colégio têm distribuição aproximadamente normal, com média 170 cm e desvio padrão 5 cm.

(a) Qual o número esperado de alunos com altura superior a 165 cm?

(b) Qual o intervalo simétrico em torno da média que conterá 75% das alturas dos alunos?

18. As vendas de determinado produto têm distribuição aproximadamente normal, com média 500 unidades e desvio padrão 50 unidades. Se a empresa decide fabricar 600 unidades no mês em estudo, qual é a probabilidade de que não possa atender a todos os pedidos desse mês, por estar com a produção esgotada?

19. Suponha que as amplitudes de vida de dois aparelhos elétricos, D_1 e D_2, tenham distribuições $N(42, 36)$ e $N(45, 9)$, respectivamente. Se os aparelhos são feitos para ser usados por um período de 45 horas, qual aparelho deve ser preferido? E se for por um período de 49 horas?

20. O diâmetro X de rolamentos esféricos produzidos por uma fábrica tem distribuição $N(0{,}6140; (0{,}0025)^2)$. O lucro T de cada rolamento depende de seu diâmetro. Assim,

$T = 0{,}10$, se o rolamento for bom $(0{,}610 < X < 0{,}618)$;

$T = 0{,}05$, se o rolamento for recuperável $(0{,}608 < X < 0{,}610)$ ou $(0{,}618 < X < 0{,}620)$;

$T = -0{,}10$, se o rolamento for defeituoso $(X < 0{,}608 \text{ ou } X > 0{,}620)$.

Calcule:

(a) as probabilidades de que os rolamentos sejam bons, recuperáveis e defeituosos.

(b) $E(T)$.

21. Suponha que um mecanismo eletrônico tenha um tempo de vida X (em 1.000 horas) que possa ser considerado uma v.a. contínua com f.d.p. $f(x) = e^{-x}$, $x > 0$. Suponha que o custo de fabricação de um item seja 2,00 reais e o preço de venda seja 5,00 reais. O fabricante garante total devolução se $X \leqslant 0{,}9$. Qual o lucro esperado por item?

22. Seja Y com distribuição binomial de parâmetros $n = 10$ e $p = 0{,}4$. Determine a aproximação normal para:

    (a) $P(3 < Y < 8)$, (b) $P(Y \geqslant 7)$, (c) $P(Y < 5)$.

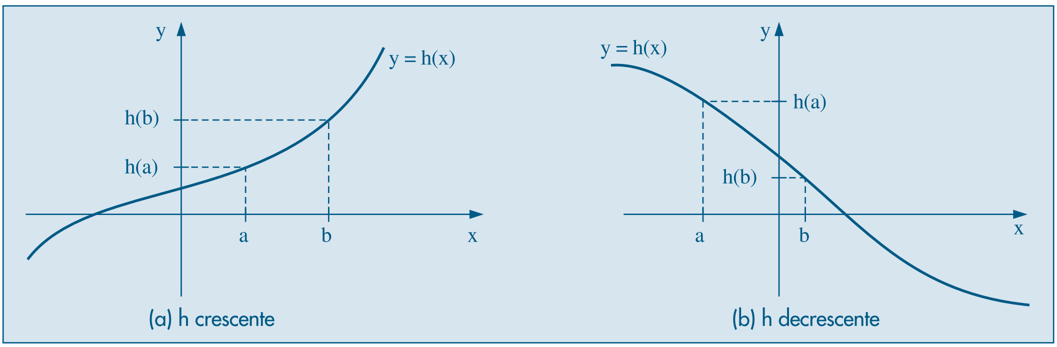
23. De um lote de produtos manufaturados, extraímos 100 itens ao acaso; se 10% dos itens do lote são defeituosos, calcule a probabilidade de 12 itens serem defeituosos. Use também a aproximação normal.

24. A confiabilidade de um mecanismo eletrônico é a probabilidade de que ele funcione sob as condições para as quais foi planejado. Uma amostra de 1.000 desses itens é escolhida ao acaso e os itens são testados, obtendo-se 30 defeituosos. Calcule a probabilidade de se obter pelo menos 30 itens defeituosos, supondo que a confiabilidade de cada item é 0,95.

7.6 Funções de Variáveis Contínuas

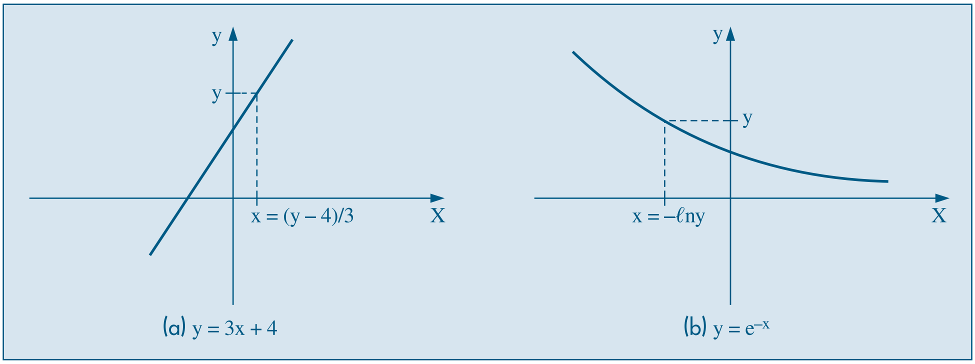
Vimos, no Capítulo 6, como obter a distribuição de uma v.a. $Y = h(X)$, se conhecermos a distribuição da v.a. discreta X. Vejamos, agora, o caso em que X é contínua. Suponhamos, primeiramente, que a função h seja estritamente monotônica, crescente ou decrescente. Neste caso, a inversa h^{-1} estará univocamente determinada e podemos obter $x = h^{-1}(y)$, para valores x e y das v.a. X e Y, respectivamente. Observando a Figura 7.22, vemos que, se a densidade de X, $f(x)$, digamos, for positiva no intervalo $a < x < b$, então a densidade de Y será positiva para $h(a) < y < h(b)$, se h for crescente, e para $h(b) < y < h(a)$, se h for decrescente.

Figura 7.22: Função de uma v.a.



Exemplo 7.11. Suponha X com a densidade do Exemplo 7.2 e considere $Y = 3X + 4$. Aqui, $y = h(x) = 3x + 4$, que é crescente (Figura 7.23 (a)).

Figura 7.23: Exemplos de funções de v.a. (a) Exemplo 7.11 (b) Exemplo 7.12.



Denotando a densidade de Y por $g(y)$, e como $f(x) > 0$ para $0 < x < 1$, $g(y) > 0$ para $4 < y < 7$.

Notemos que se podem obter probabilidades relativas a Y a partir da densidade de X. Por exemplo,

$$P(Y > 1) = P(3X + 4 > 1) = P(X > -1) = 1.$$

Vejamos como se pode obter $g(y)$. Denotemos por $G(y)$ a função de distribuição acumulada de Y. Da seção 7.3, sabemos que $G'(y) = g(y)$, para todo valor de y para o qual G for derivável. Então, temos

$$G(y) = P(Y \leqslant y) = P(3X + 4 \leqslant y) = P\left(X \leqslant \frac{y-4}{3}\right) = F\left(\frac{y-4}{3}\right),$$

onde estamos denotando por $F(\cdot)$ a função de distribuição acumulada de X. Usando a regra da cadeia para derivadas, temos

$$G'(y) = F'\left(\frac{y-4}{3}\right) \cdot \frac{1}{3} = \frac{1}{3}\, f\left(\frac{y-4}{3}\right),$$

do que decorre

$$g(y) = \begin{cases} \dfrac{2}{9}(y-4), & \text{se } 4 < y < 7 \\ 0, & \text{caso contrário.} \end{cases}$$

Exemplo 7.12. Suponha, agora, que X tenha densidade $f(x) = 3x^2/2$, $-1 < x < 1$ e que $Y = e^{-X}$. Segue-se que $h(x) = e^{-x}$ é uma função decrescente e $x = -\ell n(y)$ (Figura 7.23 (b)). Então,

$$\begin{aligned} G(y) &= P(Y \leqslant y) = P(e^{-X} \leqslant y) = P(X \geqslant -\ell n(y)) \\ &= 1 - P(X \leqslant -\ell n(y)) = 1 - F(-\ell n(y)), \end{aligned}$$

onde novamente F denota a f.d.a. de X. Derivando, obtemos a f.d.p. de Y,

$$g(y) = \frac{3}{2y} (\ell n(y))^2, \; e^{-1} < y < e.$$

O seguinte resultado generaliza esses dois exemplos.

Teorema 7.1. Se X for uma v.a. contínua, com densidade $f(x) > 0$, $a < x < b$, então $Y = h(X)$ tem densidade

$$g(y) = f(h^{-1}(y)) \left| \frac{dx}{dy} \right|, \tag{7.30}$$

supondo que h seja monotônica, derivável para todo x. Se h for crescente, $g(y) > 0$, $h(a) < y < h(b)$ e, se h for decrescente, $g(y) > 0$, $h(b) < y < h(a)$.

Prova. Basta notar que $G(y) = P(Y \leqslant y) = P(h(X) \leqslant y)$ e que essa probabilidade é igual a $P(X \leqslant h^{-1}(y)) = F(h^{-1}(y))$, se h for crescente, e igual a $1 - F(h^{-1}(y))$, se h for decrescente. Derivando $G(y)$ obtemos o resultado, notando que a derivada $(h^{-1}(y))' = dx/dy > 0$ se h for crescente, e negativa se h for decrescente.

Suponha, agora, que h não seja monotônica. Um caso de interesse que será usado mais tarde é $Y = h(X) = X^2$ (Figura 7.24). Temos

$$\begin{aligned} G(y) &= P(Y \leqslant y) = P(X^2 \leqslant y) = P(-\sqrt{y} \leqslant X \leqslant \sqrt{y}) \\ &= F(\sqrt{y}) - F(-\sqrt{y}), \end{aligned}$$

e derivando obtemos a densidade de Y,

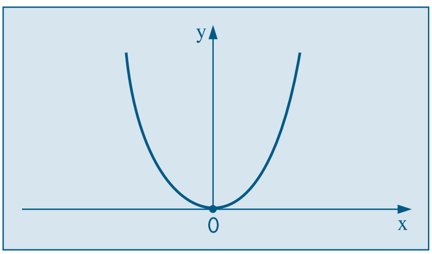
$$g(y) = \frac{1}{2\sqrt{y}} [\, f(\sqrt{y}) + f(-\sqrt{y})], \tag{7.31}$$

onde f é a densidade de X.

Se $f(x) = 1$, $0 < x < 1$ (X é uniforme no intervalo $[0, 1]$), então

$$g(y) = \frac{1}{2\sqrt{y}}, \; 0 < y < 1.$$

Figura 7.24: Ilustração de $Y = h(X) = X^2$.



Problemas

25. Considere a v.a. X do Problema 2 e $Y = X + 5$.
    (a) Calcule $P(Y \leqslant 5{,}5)$.
    (b) Obtenha a densidade de Y.
    (c) Obtenha a densidade de $Z = 2X$.

26. Suponha que a v.a. X tenha a densidade do Problema 8. Se $Y = 2X - 3/5$, obter a densidade de Y. Calcule $E(Y)$ e $\mathrm{Var}(Y)$.

27. Suponha $X \sim \mathcal{U}[-1, 1]$. Calcule a densidade de $Y = X^2$ e de $W = |X|$.

7.7 Outros Modelos Importantes

Nesta seção vamos introduzir alguns modelos para v.a. contínuas que serão bastante utilizados na terceira parte deste livro. Juntamente com o modelo normal, esses modelos são úteis para as v.a. de interesse prático, que na maioria dos casos assumem valores positivos e tendem a ter distribuições assimétricas à direita.

7.7.1 A Distribuição Gama

Uma extensão do modelo exponencial é estudado a seguir.

Definição. A v.a. contínua X, assumindo valores positivos, tem uma distribuição gama com parâmetros $\alpha > 0$ e $\beta > 0$, se sua f.d.p. for dada por

$$f(x;\, \alpha,\, \beta) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\beta^{\alpha}}\, x^{\alpha - 1} e^{-x/\beta}, & x > 0, \\ 0, & x < 0. \end{cases} \tag{7.32}$$

Em (7.32), $\Gamma(\alpha)$ é a *função gama*, importante em muitas áreas da Matemática, dada por

$$\Gamma(\alpha) = \int_0^{\infty} e^{-x} x^{\alpha - 1} dx,\ \alpha > 0. \tag{7.33}$$

Não é difícil ver que $\Gamma(\alpha) = (\alpha - 1)\,\Gamma(\alpha - 1)$, se $\alpha = n$ for um inteiro positivo, $\Gamma(n) = (n - 1)!$ e que $\Gamma(1) = 1$, $\Gamma(1/2) = \sqrt{\pi}$. Veja o Problema 45.

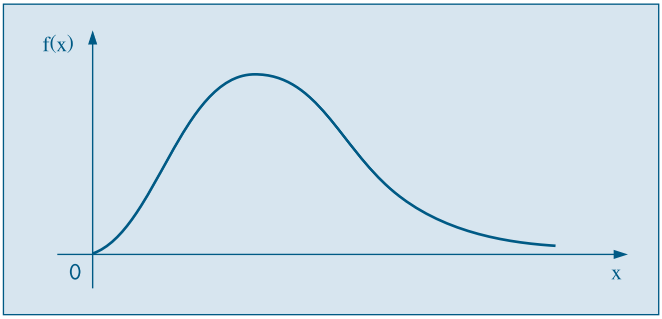
A Figura 7.25 ilustra a densidade (7.32) para $\alpha = 3$ e $\beta = 1$. Se $\alpha = 1$ obtemos a distribuição exponencial (7.26). Muitos casos de interesse têm α inteiro positivo.

Usaremos a notação

$$X \sim \mathrm{Gama}(\alpha,\, \beta)$$

para designar uma v.a. com a distribuição dada por (7.32).

Figura 7.25: Gráfico da f.d.p. de uma distribuição gama, $\alpha = 3$, $\beta = 1$.



Pode-se demonstrar que:

$$E(X) = \alpha\beta,\ \mathrm{Var}(X) = \alpha\beta^2. \tag{7.34}$$

7.7.2 A Distribuição Qui-Quadrado

Um caso especial importante do modelo gama é obtido fazendo-se $\alpha = \nu/2$ e $\beta = 2$, com $\nu > 0$ inteiro.

Definição. Uma v.a. contínua Y, com valores positivos, tem uma distribuição qui-quadrado com ν graus de liberdade (denotada $\chi^2(\nu)$), se sua densidade for dada por

$$f(y;\ \nu) = \begin{cases} \dfrac{1}{\Gamma(\nu/2)2^{\nu/2}} y^{\nu/2-1} e^{-y/2}, & y > 0 \\ 0, & y < 0. \end{cases} \tag{7.35}$$

A Figura 7.26 ilustra os gráficos de (7.35) para $\nu = 1, 2, 3$. Segue-se de (7.34) que

$$E(Y) = \nu, \quad \mathrm{Var}(Y) = 2\nu. \tag{7.36}$$

A distribuição qui-quadrado tem muitas aplicações em Estatística e, como no caso da normal, existem tabelas para obter probabilidades. A Tabela IV, fornece os valores de y_0 tais que $P(Y > y_0) = p$, para alguns valores de p e de ν. Ver Figura 7.27.

Figura 7.26: Gráficos da distribuição qui-quadrado c²(*n*).

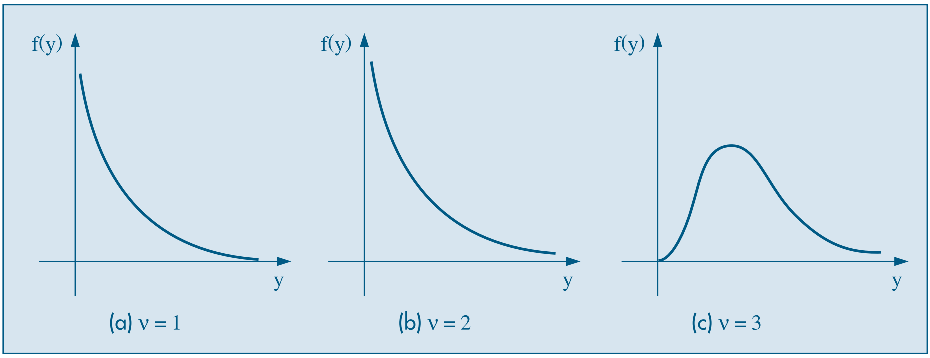
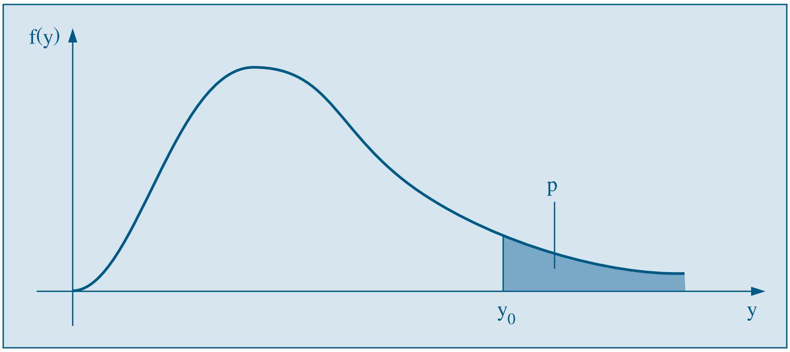


Figura 7.27: Valores tabelados da distribuição $\chi^2(\nu)$.



Exemplo 7.13. Usando a Tabela IV, para $\nu = 10$, observe que $P(Y > 2{,}558) = 0{,}99$, ao passo que $P(Y > 18{,}307) = 0{,}05$.

Para $\nu > 30$ podemos usar uma aproximação normal à distribuição qui-quadrado. Especificamente, temos o seguinte resultado: se Y tiver distribuição qui-quadrado com ν graus de liberdade, então a v.a.

$$Z = \sqrt{2Y} - \sqrt{2\nu - 1} \sim N(0{,}1).$$

Por exemplo, consultando a Tabela IV, temos que, se $\nu = 30$,

$$P(Y > 40{,}256) = 0{,}10,$$

enquanto que, usando a fórmula acima, temos que

$$z = \sqrt{2 \times 40{,}256} - \sqrt{59} = 1{,}292$$

e $P(Z > 1{,}292) = 0{,}099$, que resulta ser uma boa aproximação.

Exemplo 7.14. Considere $Z \sim N(0{,}1)$ e considere a v.a. $Y = Z^2$. De (7.31) temos que a densidade de Y é dada por

$$g(y) = \frac{1}{2\sqrt{y}} [\phi(\sqrt{y}) + \phi(-\sqrt{y})],\ y > 0,$$

onde por $\phi(z)$ indicamos a densidade da $N(0{,}1)$. Resulta

$$g(y) = \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2},$$

e comparando com (7.35) vemos que $Y \sim \chi^2(1)$. Temos, aqui, um resultado importante:

O quadrado de uma v.a. com distribuição normal padrão é uma v.a. com distribuição $\chi^2(1)$.

De um modo mais geral, uma v.a. x^2 (ν) *pode ser vista como a soma de* ν *normais padrões ao quadrado, independentes.*

7.7.3 A Distribuição *t* de Student

A distribuição t de Student é importante no que se refere a inferências sobre médias populacionais, tópico a ser tratado nos Capítulos 12 e 13. A obtenção da densidade está contida no teorema abaixo.

Teorema 7.1. Seja Z uma v.a. $N(0,1)$ e Y uma v.a. $\chi^2(\nu)$, com Z e Y independentes. Então, a v.a.

$$t = \frac{Z}{\sqrt{Y/\nu}}, \tag{7.37}$$

tem densidade dada por

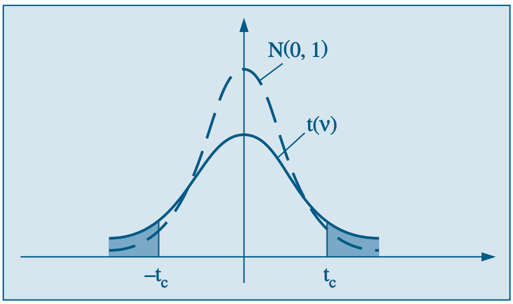
$$f(t;\,\nu) = \frac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\,(1 + t^2/\nu)^{-(\nu+1)/2},\ -\infty < t < \infty. \tag{7.38}$$

Diremos que tal variável tem uma *distribuição t de Student com* ν *graus de liberdade* e a indicaremos por $t(\nu)$. Pode-se provar que

$$E(t) = 0, \quad \text{Var}(t) = \frac{\nu}{\nu - 2},\ \nu > 2, \tag{7.39}$$

e verificar que o gráfico da densidade de t aproxima-se bastante de uma $N(0,1)$ quando ν é grande. Veja a Figura 7.28.

Figura 7.28: A distribuição t de Student e a distribuição normal padrão.



Como essa distribuição é bastante utilizada na prática, existem tabelas fornecendo probabilidades relativas a ela. A Tabela V fornece os valores de t_c tais que

$$P(-t_c < t(\nu) < t_c) = 1 - p, \tag{7.40}$$

para alguns valores de p e de ν.

O nome *Student* vem do pseudônimo usado pelo estatístico inglês W. S. Gosset, que introduziu essa distribuição no início do século passado.

Exemplo 7.15. Se $\nu = 6$, então, usando a Tabela V, $P(-1{,}943 < t(6) < 1{,}943) = 0{,}90$, ao passo que $P(t(6) > 2{,}447) = 0{,}025$. Observe que, nessa tabela, há uma linha com $\nu = \infty$, que corresponde a usar os valores da $N(0,1)$. Para $n > 120$ essa aproximação é muito boa.

7.7.4 A Distribuição *F* de Snedecor

Vamos considerar agora uma v.a. definida como o quociente de duas variáveis com distribuição qui-quadrado.

O seguinte teorema, que não será demonstrado, resume o que nos vai ser útil.

Teorema 7.2. Sejam U e V duas v.a. independentes, cada uma com distribuição qui-quadrado, com ν_1 e ν_2 graus de liberdade, respectivamente. Então, a v.a.

$$W = \frac{U/\nu_1}{V/\nu_2} \qquad (7.41)$$

tem densidade dada por

$$g(w;\ \nu_1,\ \nu_2) = \frac{\Gamma((\nu_1+\nu_2)/2)}{\Gamma(\nu_1/2)\Gamma(\nu_2/2)}\left(\frac{\nu_1}{\nu_2}\right)^{\nu_1/2} \frac{w^{(\nu_1-2)/2}}{(1+\nu_1 w/\nu_2)^{(\nu_1+\nu_2)/2}},\ w > 0. \qquad (7.42)$$

Diremos que W tem *distribuição F de Snedecor, com* ν_1 *e* ν_2 *graus de liberdade*, e usaremos a notação $W \sim F(\nu_1, \nu_2)$. Pode-se mostrar que

$$E(W) = \frac{\nu_2}{\nu_2 - 2} \quad \text{e} \quad \text{Var}(W) = \frac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2\,(\nu_2-4)}. \qquad (7.43)$$

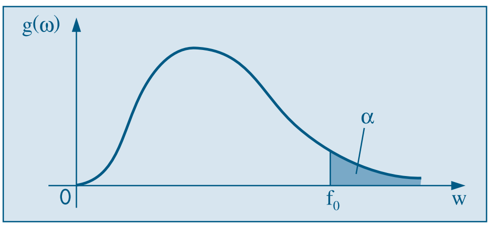
O gráfico típico de uma v.a. com distribuição F está na Figura 7.29. Na Tabela VI são dados os pontos f_0 tais que

$$P\{F(\nu_1, \nu_2) > f_0\} = \alpha,$$

para $\alpha = 0{,}05$, $\alpha = 0{,}025$ e alguns valores de ν_1 e ν_2. Para encontrar os valores inferiores, usa-se a identidade

$$F(\nu_1, \nu_2) = 1/F(\nu_2, \nu_1). \qquad (7.44)$$

Figura 7.29: Gráfico de distribuição F.



Exemplo 7.16. Considere, por exemplo, $W \sim F(5, 7)$. Consultando a Tabela VI, $P(F > 3{,}97) = 0{,}05$ ou, então, $P(F \leqslant 3{,}97) = 0{,}95$. Digamos, agora, que desejamos encontrar o valor f_0 tal que $P(F < f_0) = 0{,}05$. Da igualdade (7.44) temos

$$0{,}05 = P\{F(5{,}7) < f_0\} = P\{1/F(7{,}5) < f_0\} = P\{F(7{,}5) > 1/f_0\},$$

e procurando na Tabela VI, para $F(7{,}5)$, obtemos $1/f_0 = 4{,}88$ e, portanto, $f_0 = 0{,}205$.

Na seção de Problemas e Complementos apresentamos algumas outras distribuições de interesse, como a log-normal, Pareto, Weibull e beta.

Na Tabela 7.2 mostramos os principais modelos para v.a. contínuas, incluindo: a densidade, o domínio dos valores, os parâmetros, a média e a variância.

Tabela 7.2: Modelos para variáveis contínuas.

Modelo	$f(x)$	Parâmetros	$E(X)$, $\mathrm{Var}(X)$
Uniforme	$1/(\beta-\alpha),\ \alpha < x < \beta$	α, β	$(\alpha+\beta)/2,\ (\beta-\alpha)^2/12$
Exponencial	$1/\beta\ e^{-t/\beta},\ t > 0$	β	$\beta,\ \beta^2$
Normal	$\dfrac{1}{\sigma\sqrt{2\pi}}\ \exp\left\{\left(\dfrac{x-\mu}{\sigma}\right)^2\right\}, -\infty < x < \infty$	μ, σ	$\mu,\ \sigma^2$
Gama	$\beta^{-\alpha}/\Gamma(\alpha)\ x^{\alpha-1}\ e^{-x/\beta},\ x > 0$	$\beta > 0,\ \alpha > 0$	$\alpha\beta,\ \alpha\beta^2$
Qui-quadrado	$\dfrac{2^{-\nu/2}}{\Gamma(\nu/2)}\ y^{\nu/2-1}e^{-y/2},\ y > 0$	ν	$\nu,\ 2\nu$
t-Student	$\dfrac{\Gamma((\nu+1)/2)}{\Gamma(\nu/2)\sqrt{\pi\nu}}\left(1+\dfrac{t^2}{\nu}\right)^{-(\nu+1)/2}, -\infty < t < \infty$	ν	$0,\ \nu/(\nu-2)$
F-Snedecor	$\dfrac{\Gamma\left(\dfrac{(\nu_1+\nu_2)}{2}\right)}{\Gamma\left(\dfrac{\nu_1}{2}\right)\Gamma\left(\dfrac{\nu_2}{2}\right)}\left(\dfrac{\nu_1}{\nu_2}\right)^{\frac{\nu_1}{2}}\dfrac{w^{\frac{\nu_1-2}{2}}}{\left(1+\dfrac{\nu_1 w}{\nu_2}\right)^{\frac{\nu_1+\nu_2}{2}}},\ w > 0.$	ν_1, ν_2	$\dfrac{\nu_2}{\nu_2-2},\ \dfrac{2\nu_2^2(\nu_1+\nu_2-2)}{\nu_1(\nu_2-2)^2(\nu_2-4)}$

7.8 Quantis

No Capítulo 6 definimos o p-quantil $Q(p)$ como o valor da v.a. discreta X satisfazendo as duas desigualdades de (6.26).

No caso de uma v.a. contínua X, essa definição torna-se mais simples. Se $F(x)$ designar a f.d.a. de X, temos que as desigualdades em (6.26) ficam:

$$P(X \leqslant Q(p)) = F(Q(p)) \geqslant p \qquad (7.45)$$

e

$$P(X \geqslant Q(p)) = 1 - P(X < Q(p)) = 1 - P(X \leqslant Q(p)) = 1 - F(Q(p)) \geqslant 1 - p. \tag{7.46}$$

Mas (7.46) pode ser reescrita como

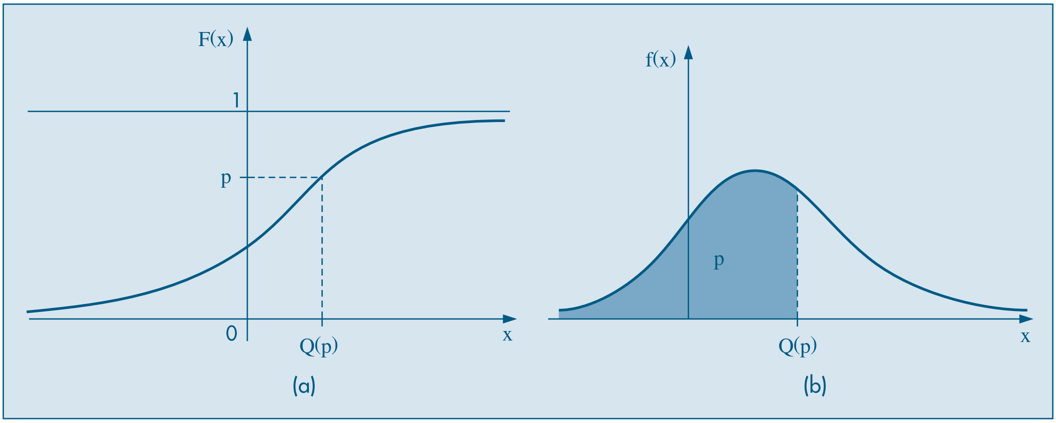
$$F(Q(p)) \leqslant p. \tag{7.47}$$

Portanto, de (7.45) e (7.47) chegamos à conclusão de que o p-quantil deve satisfazer

$$F(Q(p)) = p. \tag{7.48}$$

Graficamente, temos a situação ilustrada na Figura (7.30). Ou seja, para obter $Q(p)$, marcamos p no eixo das ordenadas, consideramos a reta horizontal pelo ponto $(0, p)$ até encontrar a curva de $F(x)$ e baixamos uma reta vertical até encontrar $Q(p)$ no eixo das abscissas. Analiticamente, temos de resolver a equação (7.48). Vejamos alguns exemplos.

Figura 7.30: Definição de $Q(p)$ (a) f.d.a. (b) f.d.p.



Exemplo 7.17. Se $Z \sim N(0, 1)$, utilizando a Tabela III encontramos facilmente que

$Q(0, 5) = Q_2 = 0$,
$Q(0, 25) = Q_1 = -0{,}675$,
$Q(0, 30) = -0{,}52$,
$Q(0{,}75) = Q_3 = 0{,}675$.

Exemplo 7.18. Suponha que $Y \sim \text{Exp}(2)$. Se quisermos calcular a mediana, Q_2, teremos de resolver

$$\int_0^{Q_2} f(y)dy = 0{,}5,$$

ou seja,

$$1/2 \int_0^{Q_2} e^{-y/2}\, dy = 0{,}5.$$

Obtemos

$$1 - e^{-Q_2/2} = 0{,}5,$$

do que temos, finalmente, $Q_2 = -2\ell n(0{,}5) = 1{,}386$.

7.9 Exemplos Computacionais

Nesta seção final, vamos dar alguns exemplos de como obter probabilidades acumuladas para a normal e exponencial, usando o pacote Minitab. Isso também pode ser feito com outros pacotes ou planilhas, bem como considerar outras distribuições contínuas.

Considere a v.a. contínua X, com f.d.a. $F(x) = P(X \leqslant x)$. O problema é, dado x, calcular $F(x)$, ou dado $F(x)$, calcular x.

Exemplo 7.19. Suponha $X \sim N(10, 25)$. Para obter $F(x)$, para $x = 8{,}65$, usamos os comandos CDF e NORMAL do Minitab. Por outro lado, se $F(x) = 0{,}8269$, então obteremos x usando os comandos INVCDF e NORMAL. Veja o Quadro 7.1.

Quadro 7.1 Obtenção de x e $F(x)$ para a Normal. Minitab.

```
MTB > CDF 8.65;
SUBC > NORMAL 10,25.
Cumulative Distribution Function
Normal with mean = 10.0000 and standard
deviation = 25.0000
      x    P(X < = x)
8.6500    0.4785
```

```
MTB > INVCDF 0.8269;
SUBC > NORMAL 10,25.
Inverse Cumulative Distribution Function
Normal with mean = 10.0000 and standard
deviation = 25.0000
P(X < = x)    x
0.8269        33.5496
```

Exemplo 7.20. O Quadro 7.2 mostra cálculos similares para distribuição exponencial, com média 0,5, ou seja, parâmetro $\beta = 2$.

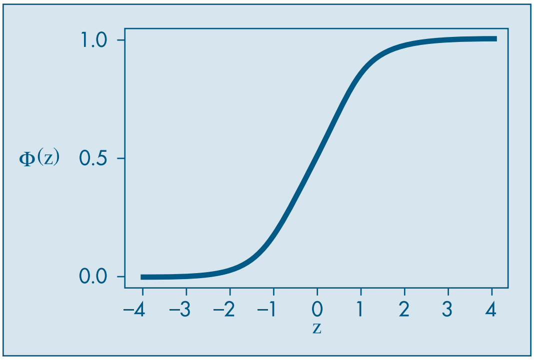
Quadro 7.2 Obtenção de x e $F(x)$ para a Exponencial. Minitab.

```
MTB > CDF 0.85;
SUBC> EXPONENCIAL 0.5.
Cumulative Distribution Function
Exponential with mean = 0.500000
      x    P(X < = x)
0.8500    0.8173
```

```
MTB > INVCDF 0.345;
SUBC> EXPONENCIAL 0.5.
Inverse Cumulative Distribution Function
Exponential with mean = 0.500000
P(X < = x)    x
0.3450        0.2116
```

Exemplo 7.21. Podemos, também, construir o gráfico de uma f.d.a, por meio de comandos do Minitab. Suponha que $Z \sim N(0,1)$. Como os valores de Z estão concentrados no intervalo [–4, 4], podemos considerar um vetor de valores z = [−4,0; −3,9; −3,8; ...; 3,8; 3,9; 4,0] e obter os valores da f.d.a. com o comando CDF. Depois, pedir para plotar os pares $(z_i, F(z_i))$. O gráfico está na Figura 7.31.

Figura 7.31: Gráfico da f.d.a. da $N(0, 1)$. Minitab.



7.10 Problemas e Complementos

28. Numa determinada localidade, a distribuição de renda (em reais) é uma v.a. X com f.d.p.

$$f(x) = \begin{cases} \dfrac{1}{10}x + \dfrac{1}{10}, & 0 \leqslant x \leqslant 2 \\ -\dfrac{3}{40}x + \dfrac{9}{20}, & 2 < x \leqslant 6 \\ 0, & x < 0 \text{ ou } x > 6. \end{cases}$$

(a) Qual a renda média nessa localidade?

(b) Escolhida uma pessoa ao acaso, qual a probabilidade de sua renda ser superior a $3.000,00?

(c) Qual a mediana da variável?

29. Se X tiver distribuição uniforme com parâmetros α e β, mostre que:

(a) $E(X) = \dfrac{\alpha + \beta}{2}$.

(b) $\text{Var}(X) = (\beta - \alpha)^2/12$.

(c) $F(x) = \begin{cases} 0, & x < \alpha \\ \dfrac{x - \alpha}{\beta - \alpha}, & \alpha \leqslant x \leqslant \beta \\ 1, & x > \beta. \end{cases}$

30. Complete a tabela abaixo, que corresponde a alguns valores da função

$$G(u) = P(0 \leqslant U \leqslant u),$$

definida na seção 7.4.1, com U uma v.a. uniforme no intervalo (–1/2, 1/2).

Probabilidades p, tais que $p = P(0 \leqslant U \leqslant u)$

Primeira decimal de u	Segunda decimal de u				Primeira decimal de u
	0	1	...	9	
0,0					0,0
0,1					0,1
0,2					0,2
0,3					0,3
0,4					0,4
0,5					0,5

31. Dada a v.a. X, uniforme em $(5, 10)$, calcule as probabilidades abaixo, usando a tabela do problema anterior.

(a) $P(X < 7)$

(b) $P(8 < X < 9)$

(c) $P(X > 8{,}5)$

(d) $P(|X - 7{,}5| > 2)$

32. Se $X \sim N(\mu, \sigma^2)$, calcular $E(X)$ e $\mathrm{Var}(X)$.

[Sugestão: Fazendo a transformação de variáveis $x = \mu + \sigma t$, obtemos que $E(X) = \frac{\mu}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-t^2/2}\,dt + \frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty} te^{-t^2/2}\,dt$. A primeira integral resulta μ (por quê?) e a segunda anula-se, pois o integrando é uma função ímpar. Para obter a variância, obtenha $E(X^2)$ por integração por partes.]

33. As notas de Estatística Econômica dos alunos de determinada universidade distribuem-se de acordo com uma distribuição normal, com média 6,4 e desvio padrão 0,8. O professor atribui graus A, B e C da seguinte forma:

Nota	Grau
$x < 5$	C
$5 \leqslant x < 7{,}5$	B
$7{,}5 \leqslant x \leqslant 10$	A

Numa classe de 80 alunos, qual o número esperado de alunos com grau A? E com grau B? E C?

34. O peso bruto de latas de conserva é uma v.a. normal, com média 1.000 g e desvio padrão 20 g.

(a) Qual a probabilidade de uma lata pesar menos de 980 g?

(b) Qual a probabilidade de uma lata pesar mais de 1.010 g?

35. A distribuição dos pesos de coelhos criados numa granja pode muito bem ser representada por uma distribuição normal, com média de 5 kg e desvio padrão de 0,8 kg. Um abatedouro comprará 5.000 coelhos e pretende classificá-los de acordo com o peso, do seguinte modo: 20% dos leves como pequenos, os 55% seguintes como médios, os 15% seguintes como grandes e os 10% mais pesados como extras. Quais os limites de peso para cada classe?

36. Uma enchedora automática de garrafas de refrigerantes está regulada para que o volume médio de líquido em cada garrafa seja de 1.000 cm^3 e o desvio padrão de 10 cm^3. Pode-se admitir que a variável volume seja normal.
   (a) Qual é a porcentagem de garrafas em que o volume de líquido é menor que 990 cm^3?
   (b) Qual é a porcentagem das garrafas em que o volume líquido não se desvia da média em mais que dois desvios padrões?
   (c) O que acontecerá com a porcentagem do item (b) se a máquina for regulada de forma que a média seja 1.200 cm^3 e o desvio padrão 20 cm^3?

37. O diâmetro de certo tipo de anel industrial é uma v.a. com distribuição normal, de média 0,10 cm e desvio padrão 0,02 cm. Se o diâmetro de um anel diferir da média em mais que 0,03 cm, ele é vendido por \$5,00; caso contrário, é vendido por \$10,00. Qual o preço médio de venda de cada anel?

38. Uma empresa produz televisores e garante a restituição da quantia paga se qualquer televisor apresentar algum defeito grave no prazo de seis meses. Ela produz televisores do tipo A (comum) e do tipo B (luxo), com lucros respectivos de \$1.000,00 e \$2.000,00, caso não haja restituição, e com prejuízos de \$3.000,00 e \$8.000,00, se houver restituição. Suponha que o tempo para a ocorrência de algum defeito grave seja, em ambos os casos, uma v.a. com distribuição normal, respectivamente, com médias 9 meses e 12 meses, e variâncias 4 meses^2 e 9 meses^2. Se tivesse de planejar uma estratégia de *marketing* para a empresa, você incentivaria as vendas dos aparelhos do tipo A ou do tipo B?

39. Determine as médias das v.a. X, Y e Z:
   (a) X uniforme em $(1, 3)$, $Y = 3X + 4$, $Z = e^X$.
   (b) X tem f.d.p. $f(x) = e^{-x}$, $x > 0$, $Y = X^2$, $Z = 3/(X + 1)^2$.

40. Suponha que X tenha distribuição uniforme em $[-a, 3a]$. Determine a média e a variância de X.

41. Se T tiver distribuição exponencial com parâmetro β, mostre que:
   (a) $E(T) = \beta$.
   (b) $\mathrm{Var}(T) = \beta^2$.

42. Os dados a seguir representam uma amostra de firmas de determinado ramo de atividade de uma região. Foram observadas duas variáveis: faturamento e número de empregados.

Nº de empregados	Nº de empresas
0 ⊢ 20	35
20 ⊢ 50	75
50 ⊢ 100	45
100 ⊢ 200	30
200 ⊢ 400	15
400 ⊢ 800	8
> 800	2
Total	210

Faturamento	Nº de empresas
0 ⊢ 10	18
10 ⊢ 50	52
50 ⊢ 100	30
100 ⊢ 200	26
200 ⊢ 400	24
400 ⊢ 800	20
800 ⊢ 1600	16
1600 ⊢ 3200	14
3200 ⊢ 6400	6
> 6400	4
Total	210

(a) Calcule a média e a variância para cada variável.

(b) Supondo normalidade para cada uma dessas variáveis, com parâmetros estimados pela amostra, calcule os valores esperados para cada intervalo de classe e compare com o observado.

43. Suponha que a v.a. X tenha densidade $f(x) = 1$, para $0 < x < 1$ e igual a zero no complementar. Faça $Y = X^2$.

(a) Determine $F_Y(y) = P(Y \leqslant y)$, y real.

(b) Determine a f.d.p. de Y.

(c) Calcule $E(X^2)$, utilizando a f.d.p. de X.

(d) Calcule $E(Y)$, utilizando a f.d.p. de Y, e compare com (c).

44. Dada a v.a.

$$Z = \frac{X - \mu_x}{\sigma_x},$$

determine a média e a variância de Z, sabendo-se que a f.d.p. de X é

$$f(x) = e^{-x},\ x > 0.$$

45. (a) Prove que, se α for inteiro positivo, $\Gamma(\alpha) = (\alpha - 1)!$.

(b) Prove que $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$.

(c) Calcule $\Gamma(1)$ e $\Gamma(1/2)$.

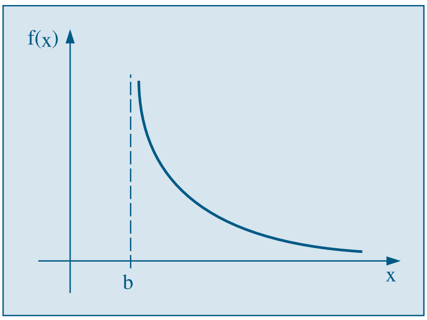
(d) Prove que a média e a variância de uma v.a. X com distribuição gama (densidade em (7.32)) são, respectivamente, $\alpha\beta$ e $\alpha\beta^2$.

46. Distribuição de Pareto. Esta é uma distribuição freqüentemente usada em Economia, em conexão com problemas de distribuição de renda.

Dizemos que a v.a. X tem *distribuição de Pareto* com parâmetros $\alpha > 0$, $b > 0$ se sua f.d.p. for dada por

$$f(x) = \begin{cases} \alpha/b\ (b/x)^{\alpha+1}, & x \geqslant b \\ 0, & x < b. \end{cases}$$

Aqui, b pode representar algum nível mínimo de renda, x é o nível de renda e $f(x)\ \Delta x$ dá a proporção de indivíduos com renda entre x e $x + \Delta x$. O gráfico de $f(x)$ está na figura abaixo.



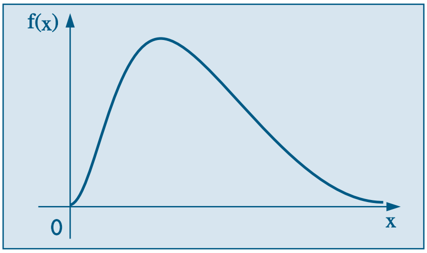
(a) Prove que $\int_{-\infty}^{\infty} f(x)dx = 1$.

(b) Mostre que, para $\alpha > 1$, $E(X) = \dfrac{\alpha b}{\alpha - 1}$ e para $\alpha > 2$, $\mathrm{Var}(X) = \dfrac{\alpha b^2}{(\alpha-1)^2(\alpha-2)}$.

47. Distribuição lognormal. Outra distribuição usada quando se têm valores positivos é a distribuição lognormal. A v.a. X tem distribuição lognormal, com parâmetros μ e σ^2, $-\infty < \mu < \infty$, $\sigma^2 > 0$, se $Y = \ell n X$ tiver distribuição normal com média μ e variância σ^2. A f.d.p. de X tem a forma

$$f(x) = \begin{cases} \dfrac{1}{x\sigma\sqrt{2\pi}}\, e^{-1/2\left(\frac{\ell n x - \mu}{\sigma}\right)^2}, & \text{se } x > 0 \\ 0, & \text{se } x \leqslant 0. \end{cases}$$

O gráfico de $f(x)$ está na figura abaixo.



(a) Prove que $E(X) = e^{\mu + \sigma^2/2}$.

(b) Se $E(X) = m$, prove que $\mathrm{Var}(X) = m^2(e^{\sigma^2} - 1)$.

48. Suponha que X tenha distribuição exponencial com parâmetro β. Prove que

$$\frac{P(X > t + x)}{P(X > x)} = P(X > t), \forall t, x \geqslant 0.$$

Essa propriedade nos diz que a distribuição exponencial não tem memória. Por exemplo, se X for a vida de um componente eletrônico, a relação acima diz que, se o componente durou até o instante x, a probabilidade de ele não falhar após o intervalo $t + x$ é a mesma de não falhar após o instante t. Nesse sentido, X "esquece" a sua idade, e a eventual falha do componente não resulta de uma deterioração gradual e sim de alguma falha repentina.

49. Se X for uma v.a. contínua, com f.d.p. $f(x)$, e se $Y = g(X)$ for uma função de X, então Y será uma v.a com

$$E(Y) = \int_{-\infty}^{\infty} g(x) f(x)\, dx.$$

Suponha que X tenha densidade

$$f(x) = \begin{cases} (½)e^{x}, & x \leqslant 0 \\ (½)e^{-x}, & x > 0. \end{cases}$$

Obtenha $E(Y)$, se $Y = |X|$.

50. Se X for uniforme no intervalo $[0, 1]$, obtenha a média da v.a. $Y = (½)X^2$.

51. Distribuição de Weibull. Um modelo que tem muitas aplicações na *teoria da confiabilidade* é o modelo de Weibull, cuja f.d.p. é dada por

$$f(x) = \begin{cases} \alpha\beta x^{\beta-1}e^{-\alpha x^\beta}, & x \geqslant 0 \\ 0, & x < 0, \end{cases}$$

onde α e β são constantes positivas. A v.a. X pode representar, por exemplo, o tempo de vida de um componente de um sistema.

(a) Se $\beta = 1$, qual a f.d.p. resultante? (b) Obtenha $E(X)$ para $\beta = 2$.

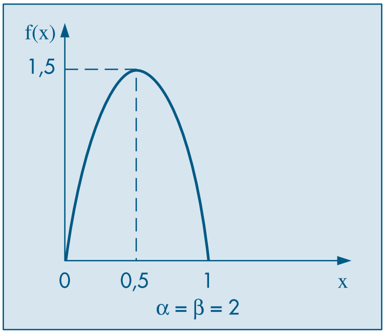
52. Distribuição Beta. Uma v.a. X tem distribuição beta com parâmetros $\alpha > 0$, $\beta > 0$, se sua f.d.p. for dada por

$$f(x) = \begin{cases} \dfrac{1}{B(\alpha, \beta)}\, x^{\alpha-1}(1-x)^{\beta-1}, & 0 < x < 1 \\ 0, & \text{caso contrário.} \end{cases}$$

Aqui, $B(\alpha, \beta)$ é a função beta, definida por

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx.$$

É possível provar que $B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$. A figura abaixo mostra a densidade da distribuição beta para $\alpha = \beta = 2$. Para esse caso, calcule $P(X \leqslant 0{,}2)$. Calcule a média e a variância de X para $\alpha = \beta = 2$.



53. Se na distribuição t de Student colocarmos $\nu = 1$, obteremos a *distribuição de Cauchy*,

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2}.$$

Mostre que $E(X)$ não existe.

54. Obtenha o gráfico da f.d.a. de uma v.a. $T \sim \text{Exp}(0, 5)$, ou seja, $E(T) = 2$, considerando 20 valores de T e calculando os valores de $F(t)$, como na seção 7.9.

55. Idem, para 30 valores de uma uniforme no intervalo $[-1,1]$.

56. Obtenha os quantis $Q(0{,}1)$, Q_1, Q_2, Q_3, $Q(0{,}9)$ para uma v.a. $X \sim N(10; 16)$.

57. Resolva a mesma questão para uma v.a. $Y \sim \chi^2(5)$.

58. Para uma v.a. com distribuição qui-quadrado, com ν graus de liberdade e ν *par*, vale a seguinte fórmula:

$$P(\chi^2(\nu) > c) = e^{-c/2} \sum_{j=0}^{\nu/2-1} \frac{(c/2)^j}{j!}.$$

Calcule essa probabilidade para os seguintes casos e compare com os valores tabelados na Tabela IV:

(a) $\nu = 4$, $c = 9{,}488$; (b) $\nu = 10$, $c = 16$.

59. Usando a aproximação normal a uma variável qui-quadrado, calcular:

(a) $P(\chi^2(35) > 49{,}76)$; (b) o valor y tal que $P(\chi^2(40) > y) = 0{,}05$.

60. Se $X \sim N(\mu\ ,\ \sigma^2)$, com densidade $f(x)$ dada por (7.17), provemos que a integral $I = \int_{-\infty}^{\infty} f(x)\, dx = 1$. Como esta integral é sempre positiva, mostremos que $I^2 = 1$. Novamente, como no Problema 32, fazemos a transformação $x = \mu + \sigma t$ e obtemos $I^2 = \frac{1}{2\pi} \int\int e^{-(t^2+s^2)/2} dsdt$, onde os limites de integração são $-\infty$ e ∞. Agora fazemos outra transformação, passando de coordenadas cartesianas para polares: $s = r \cos \theta$, $t = r \operatorname{sen} \theta$, de modo que $dsdt = r\, drd\theta$. Segue-se, integrando primeiro com relação a r e depois com relação a θ, que

$$I^2 = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} e^{-r^2/2} r dr d\theta = \frac{1}{2\pi} \int_0^{2\pi} [-e^{-r^2/2}]_0^{\infty} d\theta = \frac{1}{2\pi} \int_0^{2\pi} d\theta = 1.$$