Universidade Federal Fluminense

Instituto de Matemática e Estatística

Departamento de Estatística

Variáveis Aleatórias Contínuas

Ana Maria Lima de Farias
Departamento de Estatística

Sumário

Capítulo 1

Variáveis Aleatórias Contínuas

1.1 Noções básicas

No estudo das distribuições de frequência para variáveis quantitativas contínuas, vimos que, para resumir os dados, era necessário agrupar os valores em classes. O histograma e o polígono de frequências eram os gráficos apropriados para representar tal distribuição. Para apresentar os conceitos básicos relativos às variáveis aleatórias contínuas, vamos considerar os histogramas e respectivos polígonos de frequência apresentados na Figura 1.1. Esses gráficos representam as distribuições de frequências de um mesmo conjunto de dados, cada uma com um número de classes diferente — no histograma superior, há menos classes do que no histograma inferior. Suponhamos, também que as áreas de cada retângulo sejam iguais às frequências relativas das respectivas classes (essa é a definição mais precisa de um histograma). Por resultados vistos anteriormente, sabemos que a soma das áreas dos retângulos é 1 (as frequências relativas devem somar 1 ou 100%) e que cada frequência relativa é uma aproximação para a probabilidade de um elemento pertencer à respectiva classe.

Analisando atentamente os dois gráficos, podemos ver o seguinte: à medida que aumentamos o número de classes, diminui a diferença entre a área total dos retângulos e a área abaixo do polígono de frequência.

A divisão em classes se fez pelo simples motivo de que uma variável contínua pode assumir um número não-enumerável de valores. Faz sentido, então, pensarmos em reduzir, cada vez mais, o comprimento de classe δ, até a situação limite em que $\delta \to 0$. Nessa situação limite, o polígono de frequências se transforma em uma curva na parte positiva (ou não-negativa) do eixo vertical, tal que a área sob ela é igual a 1. Essa curva será chamada *curva densidade de probabilidade*.

Considere, agora, a Figura 1.2, onde ilustramos um fato visto anteriormente: para estimar a frequência de valores da distribuição entre os pontos a e b, podemos usar a área dos retângulos sombreados de cinza claro.

Conforme ilustrado na Figura 1.3, a diferença entre essa área e a área sob o polígono de frequências tende a diminuir, à medida que aumenta-se o número de classes. Essa diferença é a parte sombreada de cinza mais escuro. Isso nos permite concluir, intuitivamente, o seguinte: no limite, quando $\delta \to 0$, podemos estimar a probabilidade de a variável de interesse estar entre dois valores a e b pela área sob a curva densidade de probabilidade, delimitada pelos pontos a e b.

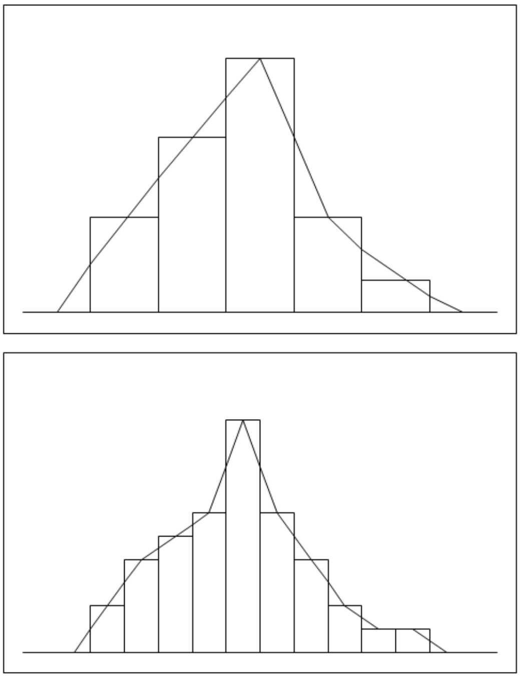


Figura 1.1 – Histograma e polígono de frequência de uma variável contínua

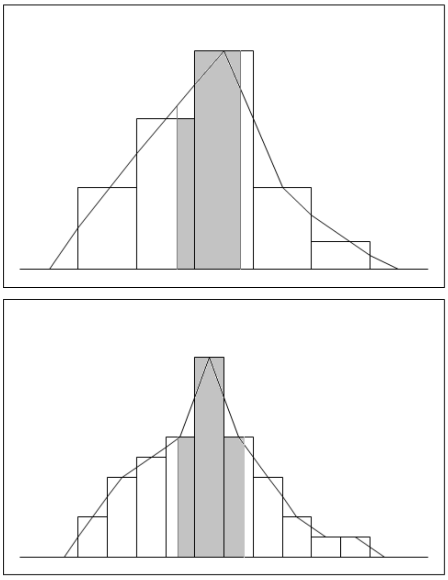


Figura 1.2 – Probabilidade como frequência relativa

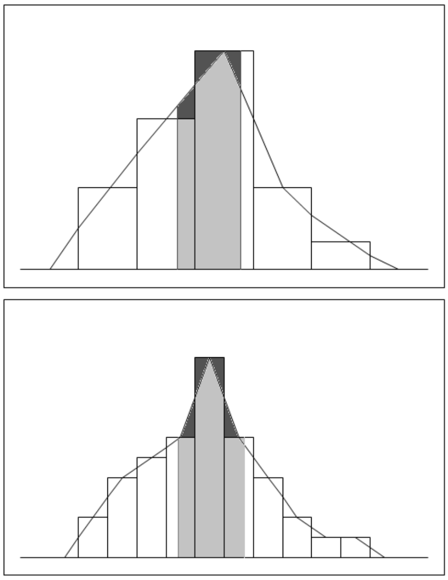


Figura 1.3 – Probabilidade como área sob o polígono de frequência

1.2 Variável aleatória contínua

Apresentamos, mais uma vez, o conceito de variável aleatória, que já foi visto no estudo das variáveis discretas, por ser este um conceito muito importante. Relembramos também as definições de variáveis aleatórias discretas e contínuas.

DEFINIÇÃO Variável aleatória

Uma **variável aleatória** é uma função real (isto é, que assume valores em $\mathbb{R}$), definida no espaço amostral Ω de um experimento aleatório. Dito de outra forma, uma variável aleatória é uma função que associa um número real a cada evento de Ω.

DEFINIÇÃO Variáveis aleatórias discretas e contínuas

Uma variável aleatória é **discreta** se sua imagem (ou conjunto de valores que ela assume) for um conjunto finito ou enumerável. Se a imagem for um conjunto não enumerável, dizemos que a variável aleatória é **contínua**.

1.3 Função densidade de probabilidade

Os valores de uma variável aleatória contínua são definidos a partir do espaço amostral de um experimento aleatório. Sendo assim, é natural o interesse na probabilidade de obtenção de diferentes valores dessa variável. O comportamento probabilístico de uma variável aleatória contínua será descrito pela sua *função densidade de probabilidade*.

Inicialmente apresentamos a definição da função densidade de probabilidade utilizando a noção de área, para seguir a apresentação inicial que considerou um histograma de uma variável contínua.

DEFINIÇÃO Função densidade de probabilidade

Uma **função densidade de probabilidade** é uma função $f(x)$ que satisfaz as seguintes propriedades:

- $f(x) \geq 0$
- A área total sob o gráfico de $f(x)$ é igual a 1.
- Dada uma função $f(x)$ satisfazendo as propriedades acima, então $f(x)$ representa alguma variável aleatória contínua X, de modo que $\mathrm{P}(a \leq X \leq b)$ é a área sob a curva limitada pelos pontos a e b (veja a Figura 1.4).

A definição acima usa argumentos geométricos; no entanto, uma definição mais precisa envolve o conceito de *integral* de uma função de uma variável, que, como se sabe, representa a área sob o gráfico da função.

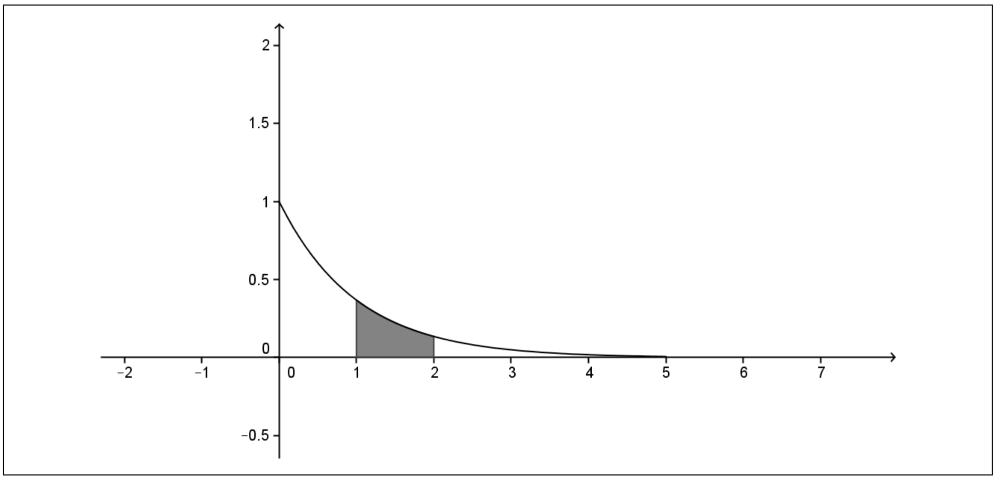


Figura 1.4 – Probabilidade como área sob a curva da função densidade de probabilidade

DEFINIÇÃO Função densidade de probabilidade

Uma **função densidade de probabilidade** é uma função $f(x)$ que satisfaz as seguintes propriedades:

- $f(x) \geq 0$
- $\int f(x)dx = 1$
- Dada uma função $f(x)$ satisfazendo as propriedades acima, então $f(x)$ representa alguma variável aleatória contínua X , de modo que

$$\mathrm{P}(a \leq X \leq b) = \int_a^b f(x)dx$$

Para deixar clara a relação entre a função densidade de probabilidade e a respectiva variável aleatória X, usaremos a notação $f_X(x)$.

Uma primeira observação importante que pode ser vista imediatamente da interpretação geométrica de probabilidade como área sob a curva densidade de probabilidade é a seguinte: se X é uma variável aleatória contínua, então a probabilidade do evento $X = a$ é zero, ou seja, a probabilidade de X ser exatamente igual a um valor específico é nula. Isso pode ser visto na Figura 1.4: o evento $\{X = a\}$ corresponde a um segmento de reta e tal segmento tem área nula. Em termos de integral, esse resultado corresponde ao fato de que $\int_a^a f(x)dx = 0$. Como consequência, temos as seguintes igualdades:

$$\mathrm{P}(a \leq X \leq b) = \mathrm{P}(a \leq X < b) = \mathrm{P}(a < X \leq b) = \mathrm{P}(a < X < b)$$

1.4 Função de distribuição

Da mesma forma que a função de probabilidade de uma variável aleatória discreta, a função densidade de probabilidade nos dá toda a informação sobre a variável aleatória contínua X, ou seja, a partir da função densidade de probabilidade, podemos calcular qualquer probabilidade associada à variável aleatória X. Também como no caso discreto, podemos calcular probabilidades associadas a uma variável aleatória contínua X a partir da f*unção de distribuição acumulada* (também denominada simplesmente função de distribuição).

DEFINIÇÃO **Função de distribuição**

Dada uma variável aleatória X, a **função de distribuição** de X é definida por

$$F_X(x) = \mathrm{P}(X \leq x) \qquad \forall x \in \mathbb{R} \tag{1.1}$$

A definição é a mesma vista para o caso discreto; a diferença é que, para variáveis contínuas, a função de distribuição acumulada é uma função contínua, sem saltos. Veja a Figura 1.5 para um exemplo.

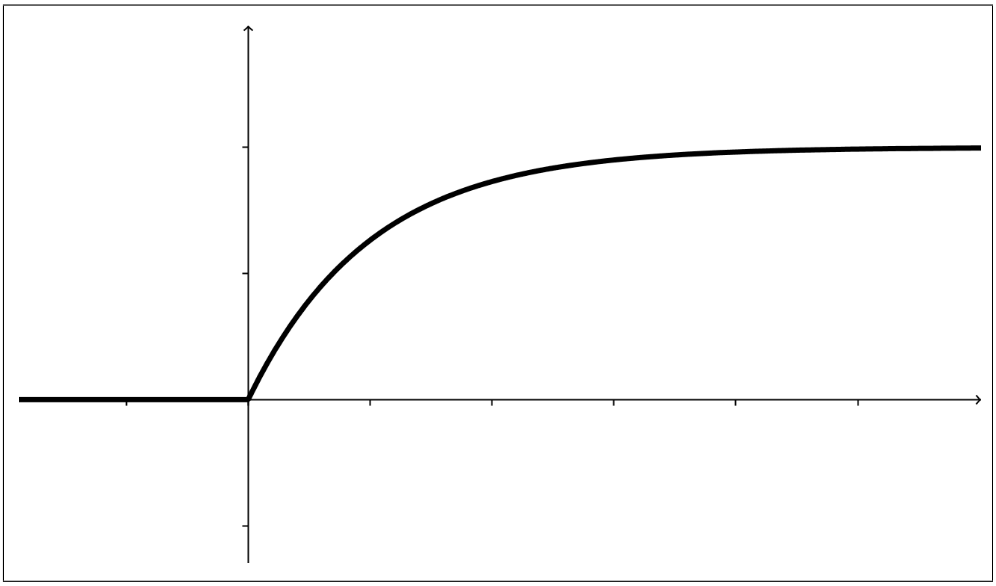


Figura 1.5 – Função de distribuição acumulada de uma v.a. contínua

Como no caso discreto, valem as seguintes propriedades para a função de distribuição acumulada de uma variável aleatória contínua:

$$0 \leq F_X(x) \leq 1 \tag{1.2}$$

$$\lim_{x \to \infty} F_X(x) = 1 \tag{1.3}$$

$$\lim_{x \to -\infty} F_X(x) = 0 \tag{1.4}$$

$$a < b \Rightarrow F_X(a) \leq F_X(b) \tag{1.5}$$

Da interpretação de probabilidade como área, resulta que $F_X(x)$ é a área à esquerda de x sob a curva densidade f_X. Veja a Figura 1.6.

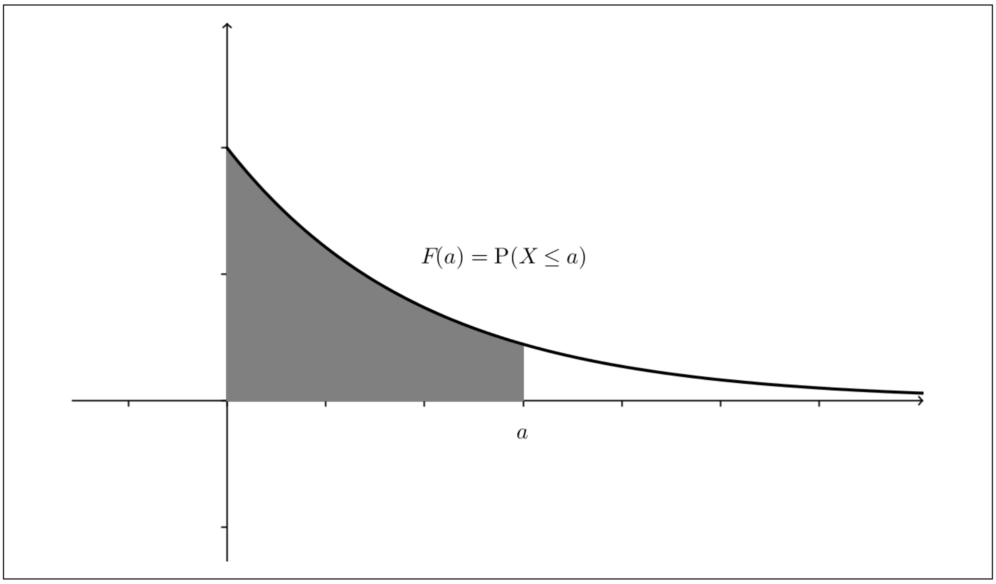


Figura 1.6 – Função de distribuição - cálculo a partir da área sob a curva densidade

Existe uma relação entre a função densidade de probabilidade e a função de distribuição acumulada, que é cosnequência do Teorema Fundamental do Cálculo.

Por definição, temos o seguinte resultado:

$$F_X(x) = \mathrm{P}(X \leq x) = \int_{-\infty}^{x} f_X(u)du \tag{1.6}$$

e do Teorema Fundamental do Cálculo resulta que

$$f_X(x) = \frac{d}{dx} F_X(x) \tag{1.7}$$

isto é, a função densidade de probabilidade é a *derivada* da função de distribuição acumulada.

1.5 Esperança e variância de variáveis aleatórias contínuas

1.5.1 Esperança

Nas distribuições de frequências agrupadas em classes de variáveis quantitativas contínuas, vimos que a média podia ser calculada como

$$\overline{x} = \sum f_i x_i$$

onde f_i era a frequência relativa da classe i e x_i era o ponto médio da classe i. Continuando com a idéia inicial de tomar classes de comprimento cada vez menor, isto é, fazendo $\delta \to 0$, chegamos à seguinte definição de esperança ou média de uma variável aleatória contínua.

DEFINIÇÃO **Esperança de uma variável aleatória contínua**

Seja X uma variável aleatória contínua com função densidade de probabilidade f_X. A **esperança** (ou **média** ou **valor esperado**) de X é definida como

$$E(X) = \int_{-\infty}^{+\infty} x f_X(x) dx \tag{1.8}$$

1.5.2 Esperança de funções de variáveis aleatórias contínuas

Se X é uma variável aleatória contínua e $h : \mathbb{R} \to \mathbb{R}$ é uma função qualquer, então $Y = h(X)$ é uma variável aleatória e sua esperança é dada por

$$E[h(X)] = \int_{-\infty}^{+\infty} h(x) f_X(x) dx \tag{1.9}$$

1.5.3 Variância

Vimos, também, que a variância, uma medida de dispersão, era calculada como a média dos desvios quadráticos em torno da média, ou seja

$$\sigma^2 = \sum f_i (x_i - \overline{x})^2$$

No caso de uma variável aleatória contínua, fazendo $h(x) = [x - E(X)]^2$, resulta novamente a definição de variância como média dos desvios quadráticos:

DEFINIÇÃO **Variância e desvio-padrão de uma variável aleatória contínua**

Seja X uma variável aleatória contínua com função de densidade de probabilidade f_X. A **variância** de X é definida como

$$\text{Var}(X) = \int_{-\infty}^{+\infty} [x - E(X)]^2 f_X(x) dx \tag{1.10}$$

O **desvio-padrão** é definido como

$$\text{DP}(X) = \sqrt{\text{Var}(X)} \tag{1.11}$$

Usando as propriedades do cálculo integral e representando por μ a esperança de X

(note que μ é uma constante, um número real), temos que:

$$\begin{aligned}
\mathrm{Var}(X) &= \int_{-\infty}^{+\infty} [x-\mu]^2 f_X(x)dx \\
&= \int_{-\infty}^{+\infty} \left(x^2 - 2\mu x + \mu^2\right) f_X(x)dx \\
&= \int_{-\infty}^{+\infty} x^2 f_X(x)dx - 2\mu \int_{-\infty}^{+\infty} x f_X(x)dx + \mu^2 \int_{-\infty}^{+\infty} f_X(x)dx
\end{aligned}$$

Se definimos $h(x) = x^2$, a primeira integral nada mais é que $\mathrm{E}(X^2)$, pelo resultado (1.9). A segunda integral é $\mathrm{E}(X) = \mu$ e a terceira integral é igual a 1, pela definição de função densidade. Logo,

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - 2\mu^2 + \mu^2 = \mathrm{E}(X^2) - \mu^2$$

o que nos leva ao resultado já visto para variáveis discretas:

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 \tag{1.12}$$

De forma resumida: a variância é a esperança do quadrado de X menos o quadrado da esperança de X.

1.5.4 Propriedades da média e da variância

As mesmas propriedades vistas para variáveis aleatórias discretas continuam valendo no caso contínuo:

Esperança	Variância	Desvio padrão
$\mathrm{E}(a) = a$	$\mathrm{Var}(a) = 0$	$\mathrm{DP}(a) = 0$
$\mathrm{E}(X+a) = \mathrm{E}(X) + a$	$\|Var(X+a) = \mathrm{Var}(X)$	$\mathrm{DP}(X+a) = \mathrm{DP}(X)$
$\mathrm{E}(bX) = b\,\mathrm{E}(X)$	$\mathrm{Var}(bX) = b^2\,\mathrm{Var}(X)$	$\mathrm{DP}(bX) = \|b\|\,\mathrm{DP}(X)$
$x_{\min} \le \mathrm{E}(X) \le x_{\max}$	$\mathrm{Var}(X) \ge 0$	$\mathrm{DP}(X) \ge 0$

Esses resultados podem ser facilmente demonstrados a partir das propriedades da integral definida e das definições vistas. Por exemplo, vamos demonstrar que $\mathrm{E}(bX) = b\,\mathrm{E}(X)$ e $\mathrm{Var}\,(bX) = b^2\,\mathrm{Var}\,(X)$. Por definição, temos que

$$\mathrm{E}(bX) = \int bxf_X(x)dx = b\int xf_X(x)dx = b\,\mathrm{E}(X)$$

Usando este resultado e a definição de variância, temos que

$$\begin{aligned}
\mathrm{Var}(bX) &= \mathrm{E}\left[bX - \mathrm{E}(bX)\right]^2 = \mathrm{E}\left\{b\left[X - \mathrm{E}(X)\right]\right\}^2 \\
&= b^2\,\mathrm{E}\left[X - \mathrm{E}(X)\right]^2 = b^2\,\mathrm{Var}(X)
\end{aligned}$$

Se interpretamos a função densidade de probabilidade de X como uma distribuição de massa na reta real, então $\mathrm{E}(X)$ é o centro de massa desta distribuição. Essa interpretação nos permite concluir, por exemplo, que se f_X é simétrica, então $\mathrm{E}(X)$ é o valor central, que define o eixo de simetria.

1.6 Exemplos

EXEMPLO 1.1 Função linear

Considere a função f_X apresentada na Figura 1.7.

(a) Encontre o valor de k para que f_X seja uma função densidade de probabilidade de uma variável aleatória X .

(b) Determine a equação que define f_X.

(c) Calcule $\Pr(2 \leq X \leq 3)$.

(d) Calcule a esperança e a variância de X.

(e) Determine o valor de k tal que $\Pr(X \leq k) = 0,6$.

(f) Encontre a função de distribuição de X.

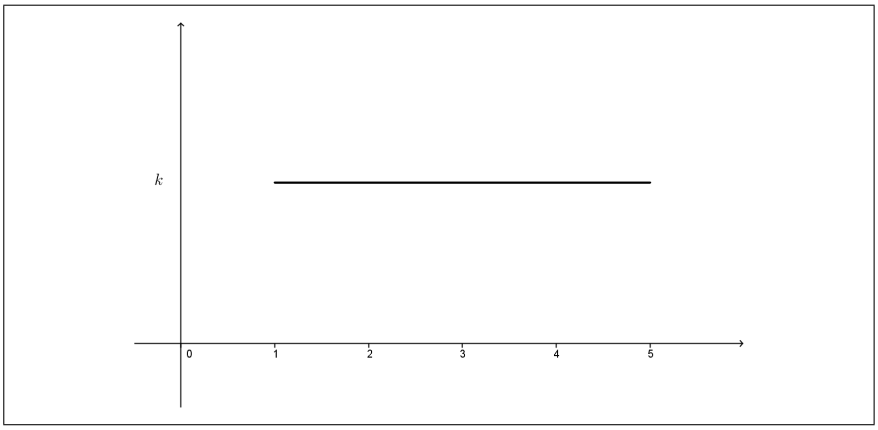


Figura 1.7 – Função densidade de probabilidade para o Exemplo 1.1

Solução

(a) A função dada corresponde a uma função constante, $f_X(x) = k$. Como a área sob a reta tem que ser 1, temos que ter

$$1 = (5-1) \times k \Rightarrow k = \frac{1}{4}$$

ou

$$\int_1^5 k dx = 1 \Rightarrow k\ x|_1^5 = 1 \Rightarrow k(5-1) = 1 \Rightarrow k = \frac{1}{4}$$

(b) Temos que

$$f_X(x) = \begin{cases} \dfrac{1}{4} & \text{se } 1 \leq x \leq 5 \\ \\ 0 & \text{caso contrário} \end{cases}$$

(c) A probabilidade pedida é a área sombreada na Figura 1.8. Logo,

$$P(2 \leq X \leq 3) = (3-2) \times \frac{1}{4} = \frac{1}{4}$$

ou

$$P(2 \leq X \leq 3) = \int_2^3 \frac{1}{4} dx = \frac{1}{4} \left. x \right|_2^3 = \frac{1}{4}$$

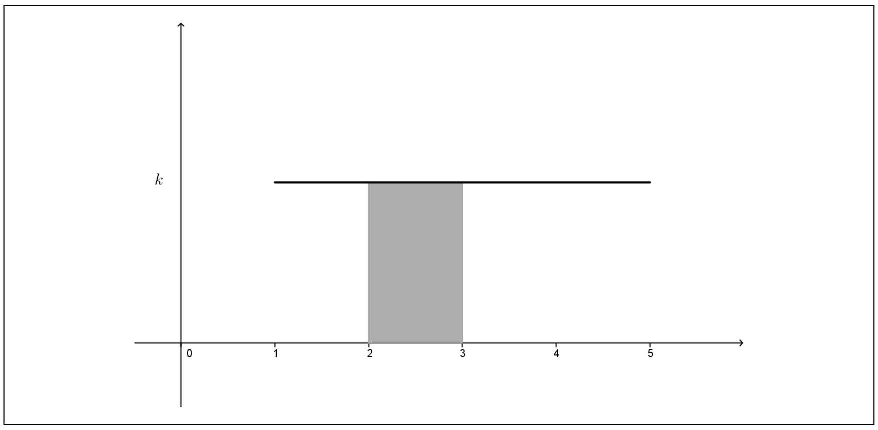


Figura 1.8 – $P(2 \leq X \leq 3)$ – Exemplo 1.1

(d) Por argumentos de simetria, a esperança é o ponto médio, ou seja, $E(X) = 3$. Usando a definição, temos:

$$E(X) = \int_1^5 x \frac{1}{4} dx = \frac{1}{4} \left(\left. \frac{x^2}{2} \right|_1^5 \right) = \frac{1}{8} (25 - 1) = 3$$

Para o cálculo da variância, temos que calcular $E(X^2)$:

$$E(X^2) = \int_1^5 x^2 \frac{1}{4} dx = \frac{1}{4} \left. \frac{x^3}{3} \right|_1^5 = \frac{1}{12} (125 - 1) = \frac{124}{12} = \frac{31}{3}$$

e

$$Var(X) = \frac{31}{3} - 3^2 = \frac{31 - 27}{3} = \frac{4}{3}$$

(e) Como a densidade é simétrica, a média e a mediana coincidem, ou seja, o ponto $x = 3$ divide a área, ou probabilidade, total ao meio. Como temos que $P(X \leq k) = 0,6$, resulta que k tem que ser maior que 3, uma vez que abaixo de 3 temos área igual a 0,5. Veja a Figura 1.9.

Temos que ter

$$0,6 = (k-1) \times \frac{1}{4} \Rightarrow k = 3,4$$

Usando integral, temos ter

$$\int_1^k \frac{1}{4} dx = 0,6 \Longrightarrow \frac{1}{4} (k-1) = 0,6 \Rightarrow k = 3,4$$

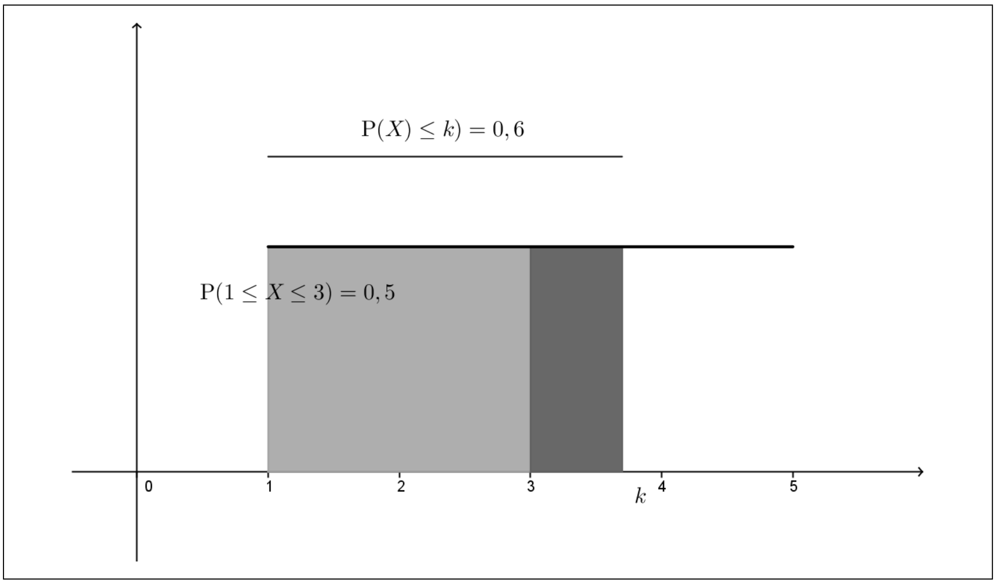


Figura 1.9 – Cálculo de k tal que $P(X \leq k) = 0,6$

(f) Para $x < 1$, temos que $F_X(x) = 0$ e para $x > 5$, temos que $F_X(x) = 1$. Para $1 \leq x \leq 5$, $F_X(x)$ é a área de um retângulo de base $(x - 1)$ e altura 1/4 (veja a Figura 1.10). Logo,

$$F_X(x) = \frac{x-1}{4}$$

e a expressão completa de F_X é

$$F_X(x) = \begin{cases} 0 & \text{se } x < 1 \\ \dfrac{x-1}{4} & \text{se } 1 \leq x \leq 5 \\ 1 & \text{se } x > 5 \end{cases}$$

cujo gráfico está ilustrado na Figura 1.11.

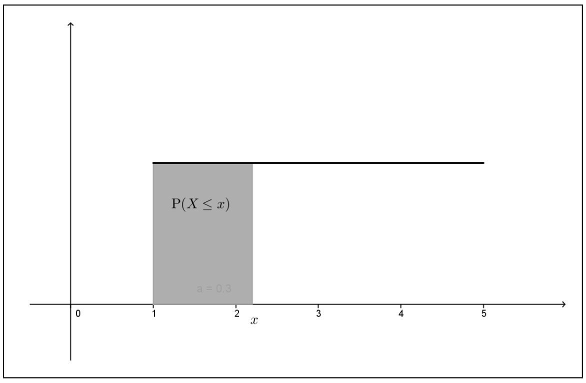


Figura 1.10 – Cálculo de $F_X(x)$ – Exemplo 1.1

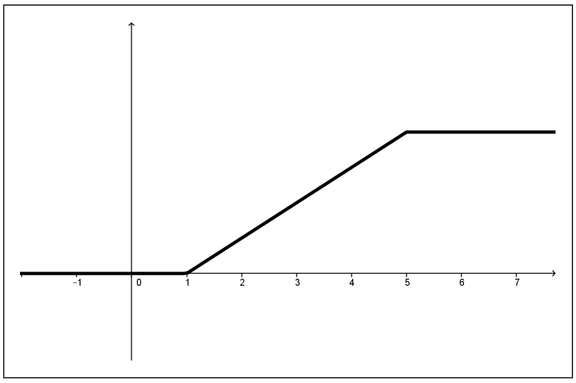


Figura 1.11 – Função de distribuição – Exemplo 1.1

♦♦

EXEMPLO 1.2 Função linear

Considere a função f_X apresentada na Figura 1.12.

(a) Encontre o valor de k para que f_X seja uma função densidade de probabilidade de uma variável aleatória contínua X e determine a equação que define f_X.

(b) Calcule $\Pr(2 \leq X \leq 3)$.

(c) Encontre a função de distribuição de X.

(d) Determine o valor de k tal que $\Pr(X \leq k) = 0,6$.

(e) Calcule a esperança e a variância de X.

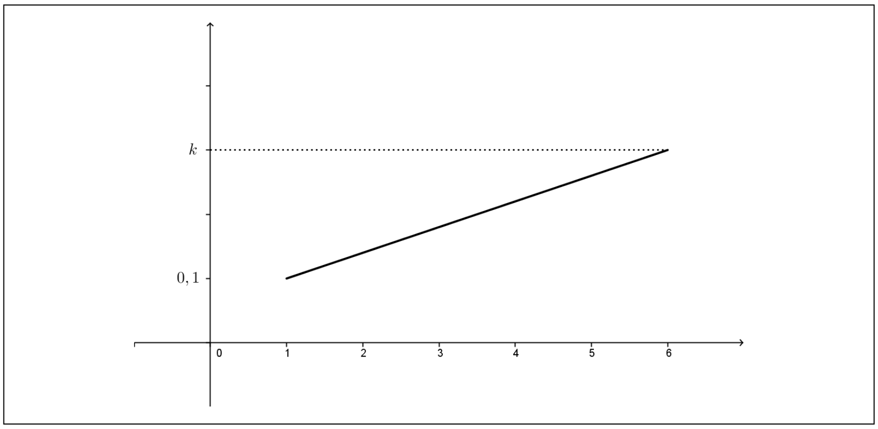


Figura 1.12 – Função densidade de probabilidade para o Exemplo 1.2

Solução

(a) Veja a Figura 1.13 – a área sob a função densidade é a área de um trapézio com bases 0,1 e k e altura 5. Como essa área tem que ser 1, resulta que

$$1 = (\frac{k + 0,1}{2} \times 5 \Rightarrow k = 0,3$$

f_X é uma função linear $f_X(x) = a + bx$ que passa pelos pontos $(1; 0,1)$ e $(6; 0,3)$, resultando, portanto, o seguinte sistema de equações:

$$\begin{cases} 0,1 = a + b \\ 0,3 = a + 6b \end{cases}$$

Subtraindo a primeira equação da segunda, obtemos

$$0,3 - 0,1 = 5b \Rightarrow b = 0,04$$

Substituindo este valor na primeira equação, obtemos que $a = 0,1 - 0,04 = 0,06$. Logo,

$$f_X(x) = \begin{cases} 0,06 + 0,04x & \text{se } 1 \leq x \leq 6 \\ 0 & \text{caso contrário} \end{cases}$$

(b) Veja a Figura 1.14, em que a área sombreada corresponde à probabilidade pedida. Vemos que essa área é a área de um trapézio de altura $3 - 2 = 1$, base maior igual a $f_X(3) = 0,06 + 0,04 \times 3 = 0,18$ e base menor igual a $f(2) = 0,06 + 0,04 \times 2 = 0,14$. Logo,

$$\Pr(2 \leq X \leq 3) = \frac{0,18 + 0,14}{2} \times 1 = 0,16$$

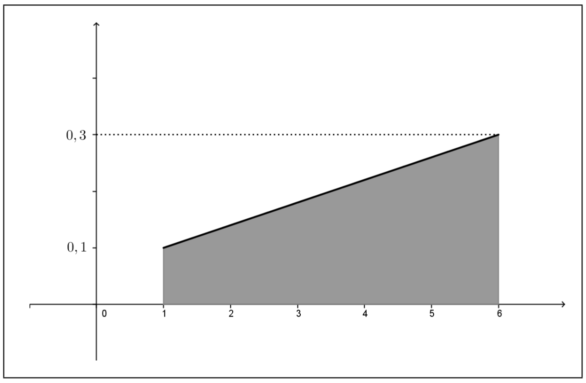


Figura 1.13 – Área sob f_X – Exemplo 1.2

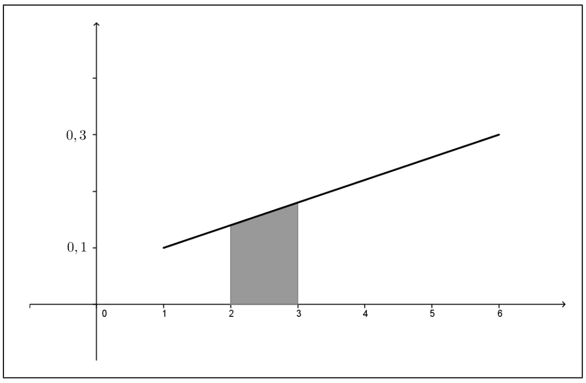


Figura 1.14 – $P(2 \leq X \leq 3)$ – Exemplo 1.2

(c) Veja a Figura 1.15; aí podemos ver que, para $x \in [1,6]$, $F_X(x)$ é a área de um trapézio de altura $x-1$; base maior igual a $f_X(x)$ e base menor igual a $f_X(1) = 0,1$. Logo,

$$\begin{aligned} F_X(x) &= \frac{(0,06+0,04x)+0,1}{2} \times (x-1) \\ &= (0,08+0,02x)(x-1) \end{aligned}$$

ou seja,

$$F_X(x) = \begin{cases} 0 & \text{se } x < 1 \\ 0,02x^2 + 0,06x - 0,08 & \text{se } 1 \leq x \leq 6 \\ 1 & \text{se } x > 6 \end{cases}$$

Na Figura 1.16 é dado o gráfico da função de distribuição.

Usando integral, temos que

$$\begin{aligned} F(x) &= \int_1^x (0,06+0,04t)dt = \left(0,06t + \frac{0,04t^2}{2}\right)\Bigg|_1^x \\ &= \left(0,06x + 0,02x^2\right) - (0,06+0,02) \\ &= 0,02x^2 + 0,06x - 0,08 \qquad 1 \leq x \leq 6 \end{aligned}$$

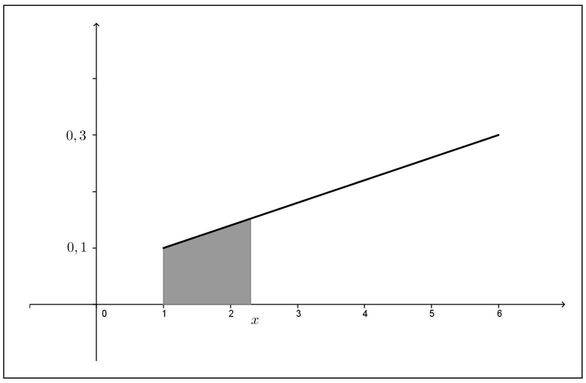


Figura 1.15 – Cálculo de $F_X(x)$ – Exemplo 1.2

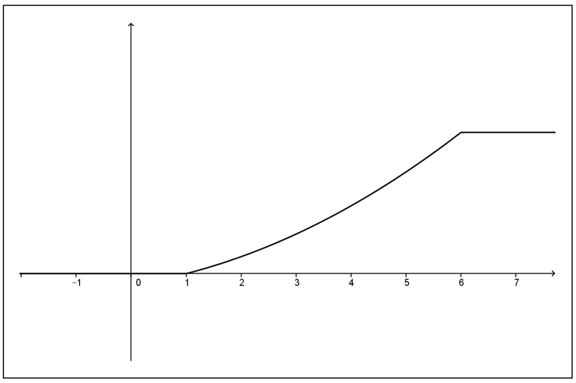


Figura 1.16 – Função de distribuição – Exemplo 1.2

(d) Queremos determinar k tal que $F_X(k) = 0,6$. Logo,

$$\begin{aligned} & 0,6 = 0,02k^2 + 0,06k - 0,08 \Rightarrow \\ & 0,02k^2 + 0,06k - 0,68 = 0 \Rightarrow \\ & k^2 + 3k - 34 = 0 \Rightarrow \\ k &= \frac{-3 \pm \sqrt{9 + 4 \times 34}}{2} \end{aligned}$$

A raiz que fornece resultado dentro do domínio de variação de X é

$$k = \frac{-3 + \sqrt{9 + 4 \times 34}}{2} \approx 4,5208$$

(e) Temos que

$$\begin{aligned}
\mathrm{E}(X) &= \int_1^6 x\,(0,06 + 0,04x)\,dx = \left(0,06\frac{x^2}{2} + 0,04\frac{x^3}{3}\right)\Big|_1^6 \\
&= \left(0,03 \cdot 36 + 0,04 \cdot \frac{6^3}{3}\right) - \left(0,03 \cdot 1 + \frac{0,04}{3}\right) \\
&= 1,08 + 2,88 - 0,03 - \frac{0,04}{3} = \frac{11,75}{3} = 3,9167
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}(X^2) &= \int_1^6 x^2\,(0,06 + 0,04x)\,dx = \left(0,06\frac{x^3}{3} + 0,04\frac{x^4}{4}\right)\Big|_1^6 \\
&= (0,02 \cdot 216 + 0,01 \cdot 1296) - (0,02 + 0,01) \\
&= 4,32 + 12,96 - 0,03 = 17,25
\end{aligned}$$

$$\mathrm{Var}(X) = 17,25 - \left(\frac{11,75}{3}\right)^2 = \frac{155,25 - 138,0625}{9} = 1,9097$$

♦♦

EXEMPLO 1.3 Densidade triangular

Considere a função f_X apresentada na Figura 1.17.

(a) Encontre o valor de h para que f_X seja uma função de densidade de probabilidade de uma variável aleatória X (note que o triângulo é isósceles!).

(b) Determine a equação que define f_X.

(c) Calcule $\Pr(1 \leq X \leq 3)$.

(d) Calcule $E(X)$ e $Var(X)$.

(e) Encontre a função de distribuição de X.

(f) Determine o valor de k tal que $\Pr(X \leq k) = 0,6$.

(g) Calcule $\mathrm{P}\left(X \leq \frac{5}{4} \middle| \frac{3}{4} \leq X \leq \frac{9}{4}\right)$.

Solução

(a) Como a área tem que ser 1, temos que ter

$$1 = \frac{1}{2} \times (4 - 0) \times h \Rightarrow h = \frac{1}{2}$$

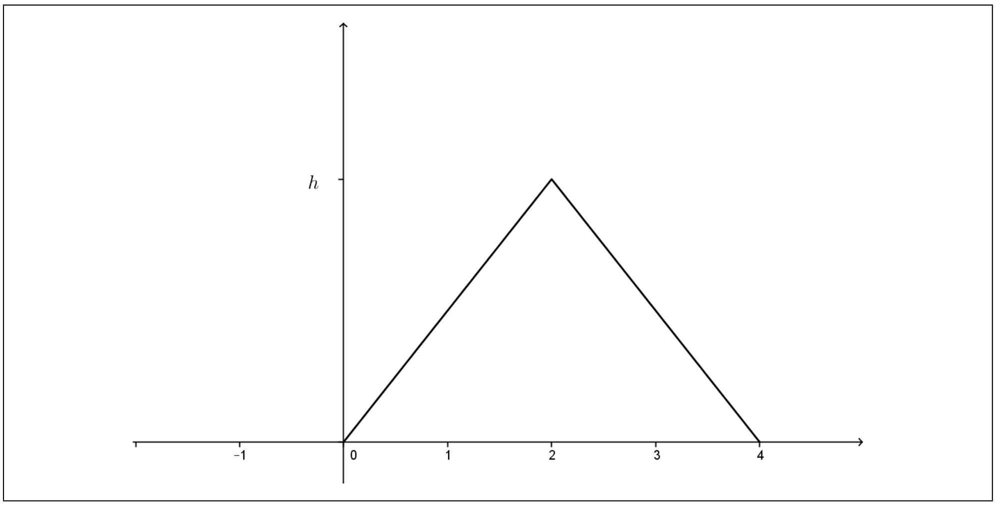


Figura 1.17 – Função densidade de probabilidade para o Exemplo 1.3

(b) A função f_X é dada por 2 equações de reta. A primeira é uma reta de inclinação positiva que passa pelos pontos $(0, 0)$ e $\left(2, \frac{1}{2}\right)$. A segunda é uma reta de inclinação negativa, que passa pelos pontos $\left(2, \frac{1}{2}\right)$ e $(4, 0)$. Para achar a equação de cada uma das retas, basta substituir as coordenadas dos dois pontos e resolver o sistema. Para a primeira reta temos o seguinte sistema:

$$\begin{aligned} 0 &= a + b \times 0 \\ \frac{1}{2} &= a + b \times 2 \end{aligned}$$

Da primeira equação resulta que $a = 0$ (é o ponto onde a reta cruza o eixo y) e substituindo esse valor de a na segunda equação, resulta que $b = \frac{1}{4}$.

Para a segunda reta, temos o seguinte sistema:

$$\begin{aligned} 0 &= a + b \times 4 \\ \frac{1}{2} &= a + b \times 2 \end{aligned}$$

Subtraindo a segunda equação da primeira, resulta

$$0 - \frac{1}{2} = (a - a) + (4b - 2b) \Rightarrow b = -\frac{1}{4}$$

Substituindo na primeira equação, encontramos que $a = 1$.

Combinando essas duas equações, obtemos a seguinte expressão para f_X :

$$f_X(x) = \begin{cases} \frac{x}{4} & \text{se } 0 \leq x < 2 \\ 1 - \frac{x}{4} & \text{se } 2 \leq x \leq 4 \\ 0 & \text{se } x < 0 \text{ ou } x > 4 \end{cases}$$

(c) A probabilidade pedida é a área sombreada em cinza claro na Figura 1.18. Os dois triângulos sombreados de cinza escuro têm a mesma área, por causa da simetria. Assim, podemos calcular a probabilidade usando a regra do complementar, uma vez que a área total é 1. A altura do triângulo inferior é $f_X(1) = \frac{1}{4}$ e a do triângulo superior é $f_X(3) = 1 - \frac{3}{4} = \frac{1}{4}$. Logo, a área de cada um dos triângulos é $\frac{1}{2} \times 1 \times \frac{1}{4} = \frac{1}{8}$ e, portanto,

$$P(1 \leq X \leq 3) = 1 - 2 \times \frac{1}{8} = \frac{6}{8} = \frac{3}{4}$$

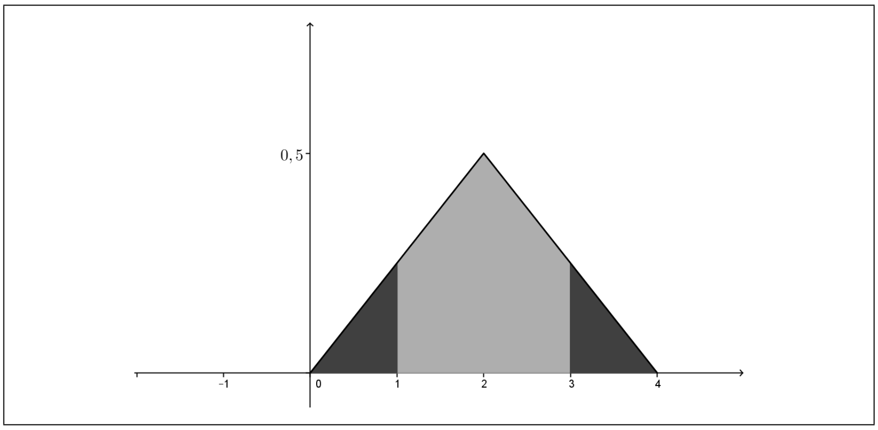


Figura 1.18 – Cálculo de $P(1 \leq X \leq 3)$ para o Exemplo 1.3

Usando integral, temos

$$\begin{aligned} P(1 \leq X \leq 3) &= \int_1^2 \frac{x}{4} dx + \int_2^3 \left(1 - \frac{x}{4}\right) dx \\ &= \frac{1}{4}\left(\frac{x^2}{2}\right)\Big|_1^2 + \left(x - \frac{x^2}{8}\right)\Big|_2^3 \\ &= \frac{1}{8}(4-1) + \left[\left(3 - \frac{9}{8}\right) - \left(2 - \frac{4}{8}\right)\right] \\ &= \frac{3}{8} + \frac{15}{8} - \frac{12}{8} = \frac{6}{8} = \frac{3}{4} \end{aligned}$$

(d) Como a função é simétrica, resulta que $E(X) = 2$.

$$\begin{aligned} E(X^2) &= \int_0^2 \frac{x^3}{4} dx + \int_2^4 x^2\left(1 - \frac{x}{4}\right) dx \\ &= \left(\frac{x^4}{16}\right)\Big|_0^2 + \left(\frac{x^3}{3} - \frac{x^4}{16}\right)\Big|_2^4 \\ &= \left(\frac{16}{16} - 0\right) + \left[\left(\frac{64}{3} - \frac{256}{16}\right) - \left(\frac{8}{3} - \frac{16}{16}\right)\right] \\ &= 1 + \frac{64}{3} - 16 - \frac{8}{3} + 1 \\ &= \frac{56}{3} - 14 = \frac{14}{3} \end{aligned}$$

$$\text{Var}(X) = \frac{14}{3} - 2^2 = \frac{2}{3}$$

(e) Assim como a função densidade de probabilidade, a função de distribuição será definida por 2 equações: uma para os valores de x no intervalo $[0,2)$ e outra para valores de x no intervalo $[2,4]$. Para $x \in [0,2)$ temos que $F_X(x)$ é a área do triângulo sombreado na Figura 1.19 e para $x \in [2,4]$, é a área sombreada na Figura 1.20 e essa área pode ser calculada pela lei do complementar.

Logo,

$$F_X(x) = \frac{1}{2}(x-0) \times \frac{x}{4} \qquad x \in [0,2)$$

Para $x \in [2, 4]$, temos que

$$F_X(x) = 1 - \frac{1}{2}(4 - x)\left(1 - \frac{x}{4}\right)$$

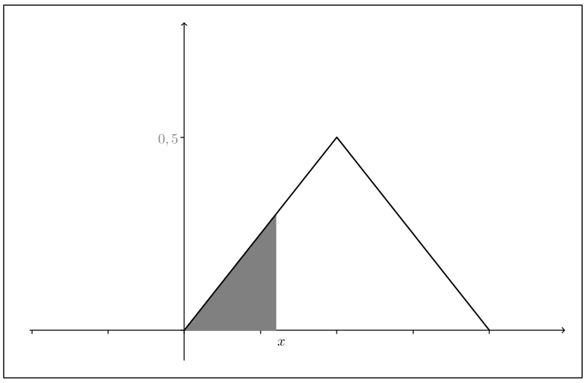


Figura 1.19 – Cálculo de $F_X(x)$ para $0 \leq x \leq 2$ – Exemplo 1.3

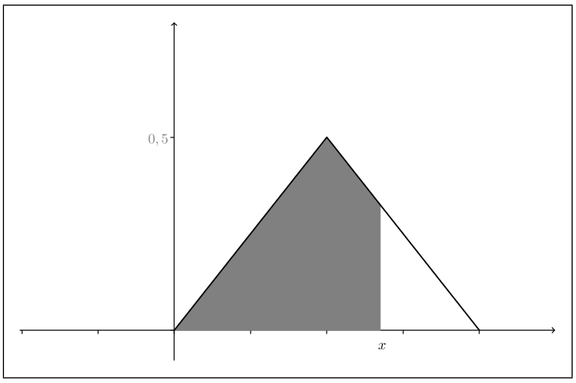


Figura 1.20 – Cálculo de $F_X(x)$ para $2 \leq x \leq 4$ – Exemplo 1.3

Combinando os resultados obtidos, resulta a seguinte expressão para F_X :

$$F_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ \frac{1}{8}x^2 & \text{se } 0 \leq x < 2 \\ 1 - \frac{1}{8}(4 - x)^2 & \text{se } 2 \leq x \leq 4 \\ 1 & \text{se } x > 4 \end{cases}$$

Veja a Figura 1.21; para $0 \leq x < 2$, o gráfico de F_X é uma parábola côncava para cima; para $2 \leq x \leq 4$, o gráfico de F_X é uma parábola côncava para baixo.

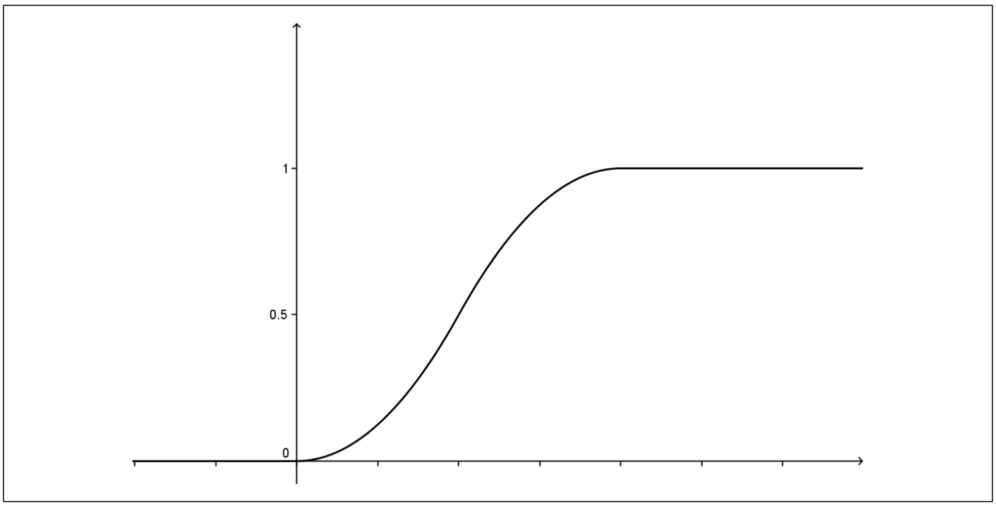


Figura 1.21 – Função de distribuição para o Exemplo 1.3

(f) Queremos determinar k tal que $F(k) = 0,6$. Como $F(2) = 0,5$, resulta que $k > 2$. Substituindo na expressão de $F(x)$, temos que ter

$$\begin{aligned}
1 - \frac{1}{8}(4-k)^2 &= 0,6 \Longrightarrow \\
1 - \frac{1}{8}\left(16 - 8k + k^2\right) &= 0,6 \Longrightarrow \\
1 - 2 + k - \frac{k^2}{8} &= 0,6 \Longrightarrow \\
\frac{k^2}{8} - k + 1,6 &= 0 \Longrightarrow \\
k^2 - 8k + 12,8 &= 0 \Longrightarrow \\
k &= \frac{8 \pm \sqrt{64 - 4 \times 12,8}}{2} = \frac{8 \pm \sqrt{12,8}}{2}
\end{aligned}$$

A raiz que fornece resultado dentro do domínio de definição de X é

$$k = \frac{8 - \sqrt{12,8}}{2} \approx 2,21$$

(g) Sabemos que $\mathrm{P}(A|B) = \dfrac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$. Assim,

$$\begin{aligned}
\mathrm{P}\left(X \leq \frac{5}{4} \middle| \frac{3}{4} \leq X \leq \frac{9}{4}\right) &= \frac{\mathrm{P}\left[\left(X \leq \frac{5}{4}\right) \cap \left(\frac{3}{4} \leq X \leq \frac{9}{4}\right)\right]}{\mathrm{P}\left(\frac{3}{4} \leq X \leq \frac{9}{4}\right)} \\
&= \frac{\mathrm{P}\left(\frac{3}{4} \leq X \leq \frac{5}{4}\right)}{\mathrm{P}\left(\frac{3}{4} \leq X \leq \frac{9}{4}\right)} = \frac{F_X\left(\frac{5}{4}\right) - F_X\left(\frac{3}{4}\right)}{F_X\left(\frac{9}{4}\right) - F_X\left(\frac{3}{4}\right)}
\end{aligned}$$

Usando a função de distribuição já calculada, temos que

$$\begin{aligned}
F_X\left(\frac{3}{4}\right) &= \frac{1}{8} \cdot \left(\frac{3}{4}\right)^2 = \frac{9}{128} \\
F_X\left(\frac{5}{4}\right) &= \frac{1}{8} \cdot \left(\frac{5}{4}\right)^2 = \frac{25}{128} \\
F_X\left(\frac{9}{4}\right) &= 1 - \frac{1}{8} \cdot \left(4 - \frac{9}{4}\right)^2 = \frac{79}{128}
\end{aligned}$$

Resulta que

$$\mathrm{P}\left(X \leq \frac{5}{4} \middle| \frac{3}{4} \leq X \leq \frac{9}{4}\right) = \frac{\frac{25}{128} - \frac{9}{128}}{\frac{79}{128} - \frac{9}{128}} = \frac{16}{70} = \frac{8}{35}$$

♦♦

EXEMPLO 1.4 Outra função linear

A variável aleatória X tem função densidade dada por

$$f(x) = \begin{cases} -x & -1 \leq x < 0 \\ x & 0 \leq x \leq 1 \\ 0 & x < -1 \text{ ou } x > 1 \end{cases}$$

(a) Esboce o gráfico de $f(x)$ e mostre que $f(x)$ realmente define uma função densidade de probabilidade.

(b) Calcule $\mathrm{P}\left(\left|X - \frac{1}{2}\right| > \frac{1}{4}\right)$.

(c) Calcule a variância de X.

(d) Obtenha a função de distribuição de X.

Solução

(a) Veja a Figura 1.22. A área de cada triângulo é igual a 1/2 e, portanto, a área total é 1. Além disso, a função é não negativa. Logo, $f(x)$ define uma função densidade de probabilidade.

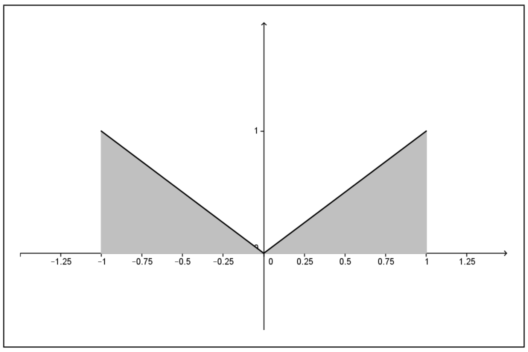


Figura 1.22 – Função densidade – Exerc. 1.4

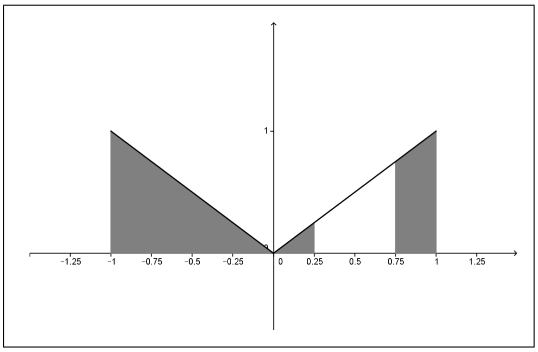


Figura 1.23 – $\mathrm{P}\left(\left|X - \frac{1}{2}\right| > \frac{1}{4}\right)$ – Exerc. 1.4

(b) Veja a Figura 1.23.

$$\begin{aligned}
\mathrm{P}\left(\left|X - \frac{1}{2}\right| > \frac{1}{4}\right) &= \mathrm{P}\left[\left(X - \frac{1}{2} > \frac{1}{4}\right) \cup \left(X - \frac{1}{2} < -\frac{1}{4}\right)\right] = \mathrm{P}\left(X > \frac{3}{4}\right) + \mathrm{P}\left(X < \frac{1}{4}\right) \\
&= \int_{3/4}^{1} x dx + 0,5 + \int_{0}^{1/4} x dx = \left[\frac{x^2}{2}\right]_{3/4}^{1} + \frac{1}{2} + \left[\frac{x^2}{2}\right]_{0}^{1/4} \\
&= \frac{1}{2} - \frac{9}{32} + \frac{1}{2} + \frac{1}{32} = 1 - \frac{1}{4} = \frac{3}{4}
\end{aligned}$$

Essa probabilidade pode, também, ser calculada por argumentos geométricos. A probabilidade pedida é 1 menos a área de um trapézio (área em branco) de base menor $0,25$, base maior $0,75$ e altura $0,5$, ou seja,

$$\mathrm{P}\left(\left|X - \frac{1}{2}\right| > \frac{1}{4}\right) = 1 - \frac{0,25 + 0,75}{2} \times 0,5 = 1 - \frac{1}{4} = \frac{3}{4}$$

(c) Como a função densidade é simétrica em torno de 0, resulta que a esperança é 0 e, nesse caso

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}(X^2) = \int_{-1}^{0} x^2(-x) dx + \int_{0}^{1} x^2 x dx = -\int_{-1}^{0} x^3 dx + \int_{0}^{1} x^3 dx \\
&= -\left[\frac{x^4}{4}\right]_{-1}^{0} + \left[\frac{x^4}{4}\right]_{0}^{1} = -\left(0 - \frac{1}{4}\right) + \left(\frac{1}{4} - 0\right) = \frac{1}{2}
\end{aligned}$$

(d) Sabemos que $F(x) = 0$ se $x < -1$ e $F(x) = 1$ se $x > 1$. Para $-1 \leq x \leq 1$, veja as Figuras 1.24 e 1.25.

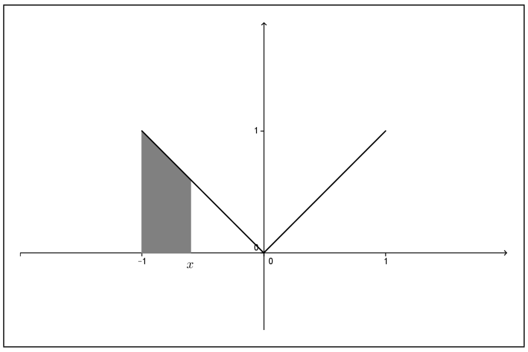


Figura 1.24 – $F(x)$, $x < 0$ – Exerc. 1.4

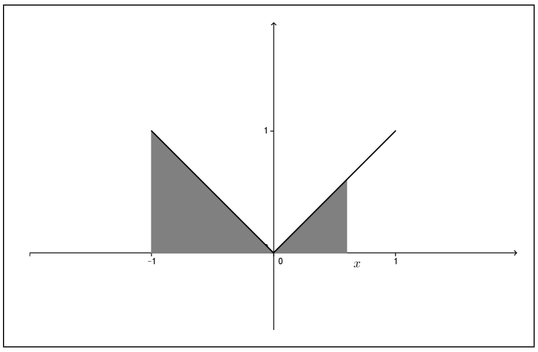


Figura 1.25 – $F(x)$, $x > 0$ – Exerc. 1.4

Para $-1 < x < 0$ temos que

$$F(x) = \int_{-1}^{x} (-t)dt = -\left[\frac{t^2}{2}\right]_{-1}^{x} = -\left(\frac{x^2}{2} - \frac{1}{2}\right) = \frac{1-x^2}{2}$$

Para $0 \leq x < 1$ temos que

$$F(x) = \int_{-1}^{0} (-t)dt + \int_{0}^{x} t dt = -\left[\frac{t^2}{2}\right]_{-1}^{0} + \left[\frac{t^2}{2}\right]_{0}^{x} = -\left(-\frac{1}{2}\right) + \left(\frac{x^2}{2} - 0\right) = \frac{1+x^2}{2}$$

Logo,

$$F(x) = \begin{cases} 0 & \text{se } x \leq -1 \\ \dfrac{1-x^2}{2} & \text{se } -1 < x < 0 \\ \dfrac{1+x^2}{2} & \text{se } 0 \leq x < 1 \\ 1 & \text{se } x \geq 1 \end{cases}$$

EXEMPLO 1.5 Demanda de um produto (Bussab & Morettin)

A demanda diária de arroz num supermercado, em centenas de quilos, é uma variável aleatória com função densidade de probabilidade

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

(a) Qual é a probabilidade de se vender mais de 150 kg num dia escolhido ao acaso?

(b) Qual a quantidade de arroz que deve ser deixada à disposição dos clientes diariamente para que não falte arroz em 95% dos dias?

Solução

(a) Seja X a variável aleatória que representa a demanda diária de arroz, em centenas de quilos. Veja a Figura 1.26, em que a área sombreada corresponde à probabilidade pedida. Nesse triângulo, a base é $3 - 1,5 = 1,5$ e a altura é $f(1,5) = \frac{-1,5}{3} + 1$. Logo,

$$\Pr(X \geq 1,5) = \frac{1}{2} \times 1,5 \times 0,5 = \frac{1}{2} \times \frac{3}{2} \times \frac{1}{2} = \frac{3}{8} = 0,375$$

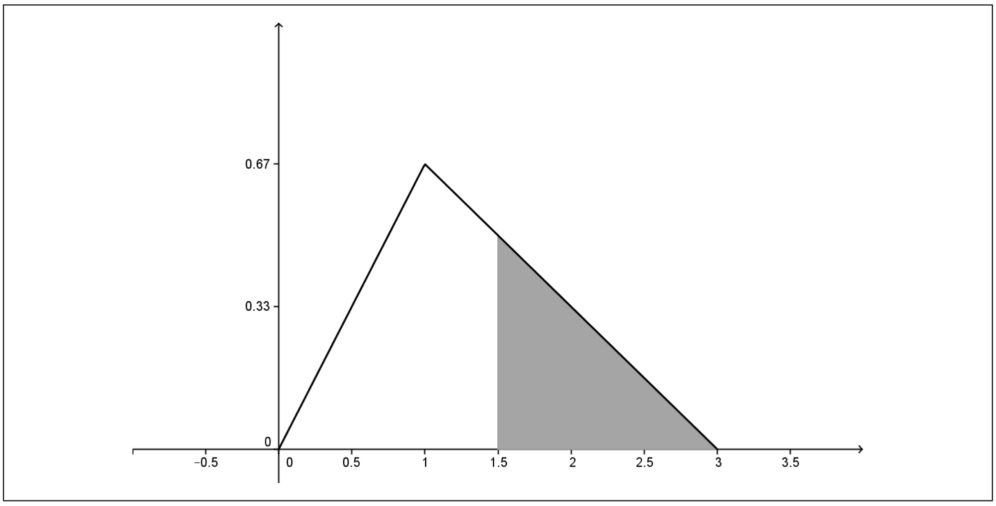


Figura 1.26 – Cálculo de $P(X > 1,5)$ – Exemplo 1.5

Em termos de integral, temos que

$$\begin{aligned} \Pr(X \geq 1,5) &= \int_{1,5}^{3} \left(-\frac{x}{3} + 1\right) dx = \left(-\frac{x^2}{6} + x\right)\Big|_{1,5}^{3} = \left(-\frac{3^2}{6} + 3\right) - \left(-\frac{1,5^2}{6} + 1,5\right) \\ &= \frac{9}{6} - \frac{6,75}{6} = \frac{2,25}{6} = 0,375 \end{aligned}$$

(b) Seja k o valor a estocar. Para que a demanda seja atendida, é necessário que a quantidade demandada seja menor que a quantidade em estoque. Logo, queremos encontrar o valor de k tal que $\Pr(X \leq k) = 0,95$.

Como $\Pr(X \leq 1) = \frac{1}{3}$, k tem que ser maior que 1, ou seja, k está no triângulo superior (veja a Figura 1.27).

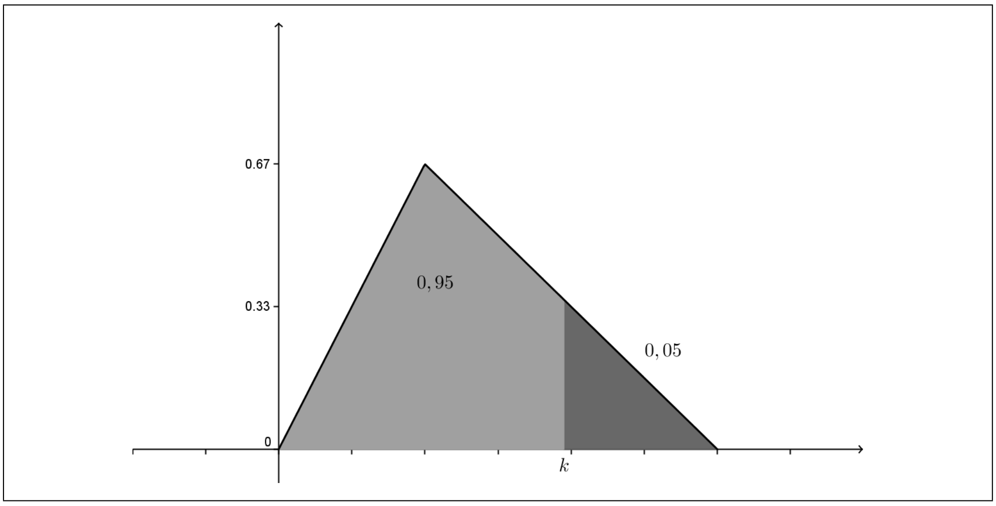


Figura 1.27 – Cálculo de k tal que $P(X \leq k) = 0,95$ – Exemplo 1.5

Mas $P(X \leq k) = 0,95$ é equivalente a $P(X > k) = 0,05$. Logo,

$$\begin{aligned}
0,05 &= \frac{1}{2}(3-k)\left(-\frac{k}{3}+1\right) \Rightarrow \\
0,1 &= (3-k)\left(\frac{-k+3}{3}\right) \Rightarrow \\
0,3 &= 9-6k+k^2 \Rightarrow \\
& k^2-6k+8,7=0 \Rightarrow \\
k &= \frac{6 \pm \sqrt{36-4\times 8.7}}{2}
\end{aligned}$$

A raiz que dá a solução dentro do domínio de X é

$$k = \frac{6-\sqrt{36-4\times 8.7}}{2} = 2,45 \text{ centenas de quilos}$$

Usando integração:

$$\begin{aligned}
P(X > k) &= 0,05 \Rightarrow \int_k^3 \left(-\frac{x}{3}+1\right) dx = 0,05 \Rightarrow \left(-\frac{x^2}{6}+x\right)\Big|_k^3 = 0,05 \Rightarrow \\
\left(-\frac{3^2}{6}+3\right) - \left(-\frac{k^2}{6}+k\right) &= 0,05 \Rightarrow \frac{k^2}{6} - k + \frac{9}{6} - 0,05 = 0 \Rightarrow k^2-6k+8,7=0
\end{aligned}$$

mesma equação obtida anteriormente.

♦♦

EXEMPLO 1.6 Sobre função de distribuição

A função de distribuição de uma variável aleatória X é dada por

$$\begin{cases} 0 & \text{se } x < 0 \\ \dfrac{3x^2 - x^3}{2} & \text{se } 0 \leq x < 1 \\ 1 & \text{se } x \geq 1 \end{cases}$$

1. Calcule $\mathrm{E}(X)$ e $\mathrm{Var}(X)$.
2. Calcule $\mathrm{P}\left(0 \leq X \leq \frac{1}{4} | X \leq \frac{1}{2}\right)$.

Solução

(a) A função densidade é a derivada da função de distribuição acumulada. Logo,

$$f(x) = \begin{cases} 3x - \dfrac{3x^2}{2} & \text{se } 0 \leq x \leq 1 \\ 0 & \text{c.c.} \end{cases}$$

$$\mathrm{E}(X) = \int_0^1 x\left(3x - \frac{3x^2}{2}\right)dx = \int_0^1 \left(3x^2 - \frac{3}{2}x^3\right)dx = \left[x^3 - \frac{3}{2}\frac{x^4}{4}\right]_0^1 = 1 - \frac{3}{8} = \frac{5}{8}$$

$$\mathrm{E}(X^2) = \int_0^1 x^2\left(3x - \frac{3x^2}{2}\right)dx = \int_0^1 \left(3x^3 - \frac{3}{2}x^4\right)dx = \left[3\frac{x^4}{4} - \frac{3}{2}\frac{x^5}{5}\right]_0^1 = \frac{3}{4} - \frac{3}{10} = \frac{9}{20}$$

$$\mathrm{Var}(X) = \frac{9}{20} - \left(\frac{5}{8}\right)^2 = \frac{9}{20} - \frac{25}{64} = \frac{144 - 125}{320} = \frac{19}{320}$$

(b)

$$\begin{aligned} \mathrm{P}\left(0 \leq X \leq \frac{1}{4}\middle| X \leq \frac{1}{2}\right) &= \frac{\mathrm{P}\left(0 \leq X \leq \frac{1}{4}\right)}{\mathrm{P}\left(X \leq \frac{1}{2}\right)} = \frac{F\left(\frac{1}{4}\right) - F(0)}{F\left(\frac{1}{2}\right)} = \frac{\frac{3\times\left(\frac{1}{4}\right)^2 - \left(\frac{1}{4}\right)^3}{2}}{\frac{3\times\left(\frac{1}{2}\right)^2 - \left(\frac{1}{2}\right)^3}{2}} \\ &= \frac{\frac{3}{16} - \frac{1}{64}}{\frac{3}{4} - \frac{1}{8}} = \frac{\frac{12-1}{64}}{\frac{6-1}{8}} = \frac{11}{64} \times \frac{8}{5} = \frac{11}{40} \end{aligned}$$

1.7 Exercícios propostos

1.1 Considere a seguinte função:

$$g(x) = \begin{cases} K(2-x) & \text{se } 0 \leq x \leq 1 \\ 0 & \text{se } x < 0 \text{ ou } x > 1 \end{cases}$$

(a) Esboce o gráfico de $g(x)$.

(b) Encontre o valor de K para que $g(x)$ seja função densidade de probabilidade de uma variável aleatória X.

(c) Encontre a função de distribuição acumulada.

(d) Calcule os quartis da distribuição.

(e) Calcule a esperança e a variância de X.

1.2 Seja X uma variável aleatória com função densidade de probabilidade dada por

$$f_X(x) = \begin{cases} 2x & \text{se } 0 \leq x \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

Calcule $\mathrm{P}\left(X \leq \frac{1}{2} \middle| \frac{1}{3} \leq X \leq \frac{2}{3}\right)$

1.3 O diâmetro de um cabo elétrico é uma variável aleatória contínua com função densidade dada por

$$f(x) = \begin{cases} k(2x - x^2) & \text{se } 0 \leq x \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

(a) Determine o valor de k.(Resp.: $k = 3/2$)

(b) Calcule $\mathrm{E}(X)$ e $\mathrm{Var}(X)$. (Resp.: $5/8; 19/320$)

(c) Calcule $\mathrm{P}(0 \leq X \leq 1/2)$. (Resp.: $5/16$)

1.4 A função densidade de probabilidade f de uma variável aleatória X é dada pela função cujo gráfico se encontra na Figura 1.28.

(a) Encontre a expressão de f.

(b) Calcule $\Pr(X > 2)$. (Resp: $1/4$)

(c) Determine m tal que $\Pr(X > m) = 1/8$. (Resp.: $m = 4 - \sqrt{2}$)

(d) Calcule a esperança e a variância de X. (Resp.: $E(X) = 4/3$; $E(X^2) = 8/3$)

(e) Calcule a função de distribuição acumulada e esboce seu gráfico.

1.5 Uma variável aleatória X tem função densidade de probabilidade dada por

$$f(x) = \begin{cases} 6x(1 - x) & \text{se } 0 \leq x \leq 1 \\ 0 & \text{caso contrário} \end{cases}$$

Se $\mu = E(X)$ e $\sigma^2 = Var(X)$, calcule $\Pr(\mu - 2\sigma < X < \mu + 2\sigma)$. (Resp.: $0,9793$)

1.6 Uma variável aleatória X tem função de distribuição acumulada F dada por

$$F(x) = \begin{cases} 0 & \text{se } x \leq 0 \\ x^5 & \text{se } 0 < x < 1 \\ 1 & \text{se } x \geq 1 \end{cases}$$

Calcule $E(X)$ e $Var(x)$. (Resp.: $5/6; 5/252$)

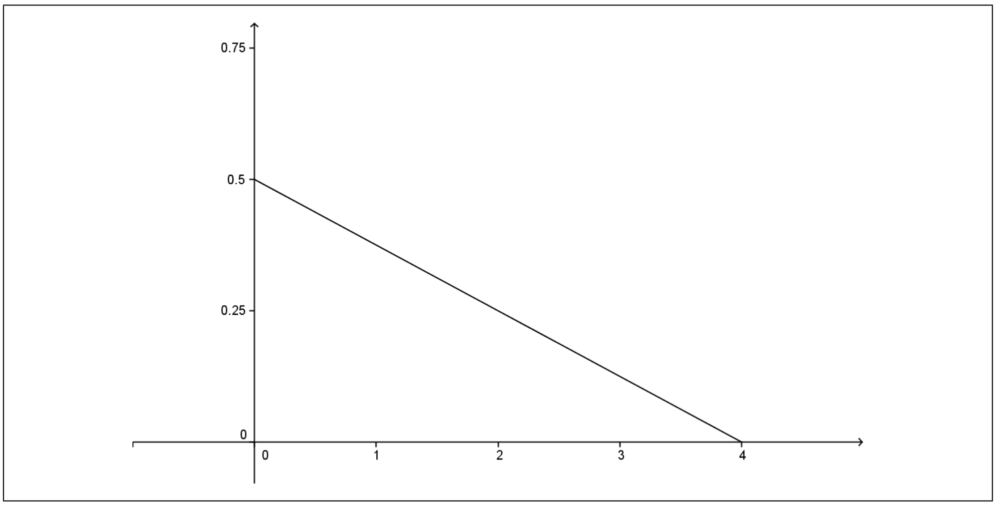


Figura 1.28 – Função densidade de probabilidade – Exercício 1.4

Capítulo 2

Algumas distribuições contínuas

2.1 Distribuição uniforme

2.1.1 Função densidade

Considere a função f_X apresentada na Figura 2.1, em que a e b são números conhecidos.

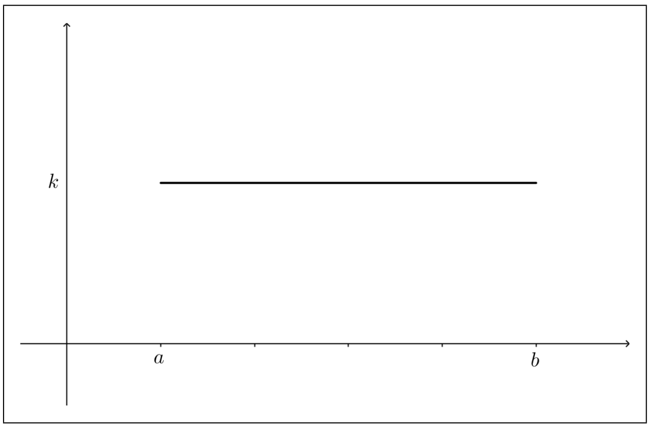


Figura 2.1 – Função densidade uniforme

Qual deve ser o valor de k para que f_X seja uma função de densidade de probabilidade de uma v.a. X?

A primeira condição é que k deve ser maior que zero e como a área tem que ser 1, resulta

$$1 = (b - a) \times k \Rightarrow k = \frac{1}{b - a}$$

Note que, para dois subintervalos de mesmo comprimento, a área será igual, uma vez que temos áreas de retângulos com mesma altura. Assim, intervalos de mesmo comprimento têm a mesma probabilidade. Esse fato leva à denominação de tal densidade como *densidade uniforme*.

DEFINIÇÃO **Distribuição uniforme**

Uma v.a. contínua X tem distribuição uniforme no intervalo $[a, b]$ se sua função densidade é dada por

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{se } x \in [a, b] \\ 0 & \text{caso contrário} \end{cases} \tag{2.1}$$

Os valores a e b são chamados *parâmetros* da distribuição uniforme; note que ambos têm de ser finitos para que a integral seja igual a 1. Quando $a = 0$ e $b = 1$ temos a uniforme padrão, denotada por $\mathcal{U}(0, 1)$.

2.1.2 Função de distribuição

Por definição, a função de distribuição acumulada é

$$F_X(x) = P(X \leq x)$$

e essa probabilidade é dada pela área sob a curva densidade à esquerda de x, conforme ilustrado na Figura 2.2.

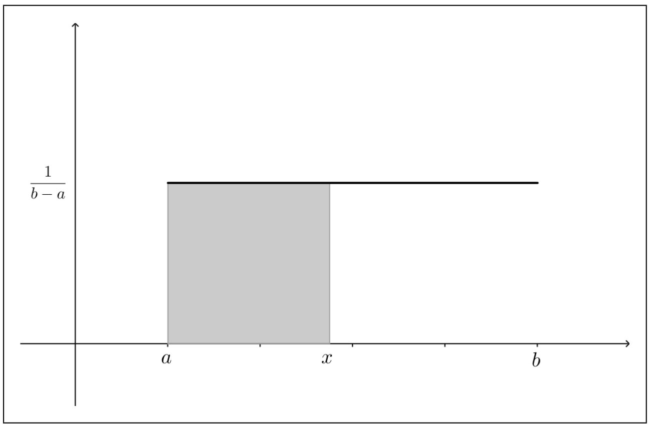


Figura 2.2 – Função de distribuição da densidade uniforme como área

Essa é a área de um retângulo com base $(x - a)$ e altura $\dfrac{1}{b-a}$. Logo,

$$F_X(x) = \begin{cases} 0 & \text{se } x < a \\ \dfrac{x-a}{b-a} & \text{se } a \leq x \leq b \\ 1 & \text{se } x > b \end{cases} \tag{2.2}$$

O gráfico dessa função de distribuição acumulada é apresentado na Figura 2.3.

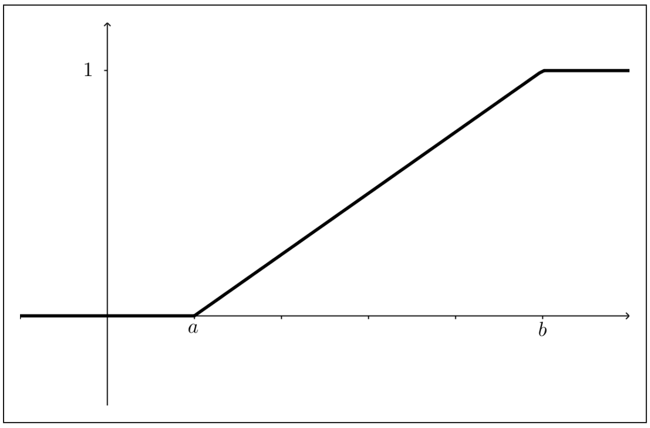


Figura 2.3 – Função de distribuição da densidade uniforme

2.1.3 Esperança e variância

Das propriedades da esperança e das características da densidade uniforme, sabemos que $E(X)$ é o ponto médio do intervalo $[a, b]$:

$$\mathrm{E}\,(X) = \frac{a+b}{2}$$

Vamos, agora, calcular $\mathrm{E}(X^2)$.

$$\mathrm{E}(X^2) = \int_a^b x^2 f(x)dx = \int_a^b x^2 \frac{1}{b-a} dx = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + b^2 + ab}{3}$$

Logo,

$$\begin{aligned} \mathrm{Var}(X) &= \frac{a^2+b^2+ab}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{4a^2+4b^2+4ab-3a^2-6ab-3b^2}{12} \\ &= \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12} \end{aligned}$$

Resumindo:

$$X \sim Unif(a, b) \quad \Rightarrow \quad \begin{cases} \mathrm{E}(X) = \dfrac{a+b}{2} \\ \\ \mathrm{Var}(X) = \dfrac{(b-a)^2}{12} \end{cases} \tag{2.3}$$

2.1.4 Exemplos

EXEMPLO 2.1 Latas de coca-cola

Latas de coca-cola são enchidas num processo automático segundo uma distribuição uniforme no intervalo (em ml) [345,355].

(a) Qual é a probabilidade de uma lata conter mais de 353 ml?

(b) Qual é a probabilidade de uma lata conter menos de 346 ml?

(c) Qualquer lata com volume 4 ml abaixo da média pode gerar reclamação do consumidor e com volume 4 ml acima da média pode transbordar no momento de abertura, devido à pressão interna. Qual é a proporção de latas problemáticas?

Solução

Seja X = "conteúdo da lata de coca-cola". Então, $X \sim U[345, 355]$

(a) Pede-se

$$\mathrm{P}(X > 353) = \frac{355 - 353}{355 - 345} = 0,2$$

(b) Pede-se

$$\mathrm{P}(X < 346) = \frac{346 - 345}{355 - 345} = 0,1$$

(c) Pede-se

$$\mathrm{P}(X < 350 - 4) + \mathrm{P}(X > 350 + 4) = \frac{346 - 345}{355 - 345} + \frac{355 - 354}{355 - 345} = 0,2$$

Logo, a proporção de latas problemáticas é de 20%. Note que essa é uma proporção bastante alta!

♦♦

EXEMPLO 2.2 Determinação dos parâmetros

Uma variável aleatória X tem distribuição uniforme no intervalo $[a, b]$ com média 7,5 e variância 6,75. Determine os valores de a e b, sabendo que $b > a > 0$.

Solução

$$\begin{cases} \dfrac{a+b}{2} = 7,5 \\ \\ \dfrac{(b-a)^2}{12} = 6,75 \end{cases}$$

Da segunda equação, resulta que $(b - a)^2 = 81$ e, portanto, $|b - a| = \sqrt{81}$. Como estamos supondo que $a < b$, resulta que $b - a = 9$, o que nos leva ao sistema

$$\begin{cases} a + b = 15 \\ b - a = 9 \end{cases}$$

cuja solução é $a = 3$ e $b = 12$, ou seja, $X \sim Unif(3, 12)$.

♦♦

2.2 Distribuição exponencial

Consideremos o gráfico da função exponencial $f(x) = e^x$, dado na Figura 2.4. Podemos ver aí que, se $x < 0$, então a área sob a curva é limitada, o mesmo valendo para uma função mais geral $f(x) = e^{\lambda x}$. Sendo assim, é possível definir uma função densidade a partir da função exponencial $e^{\lambda x}$, desde que nos limitemos ao domínio dos números reais negativos. Mas isso é equivalente a trabalhar com a função $e^{-\lambda x}$ para x positivo.

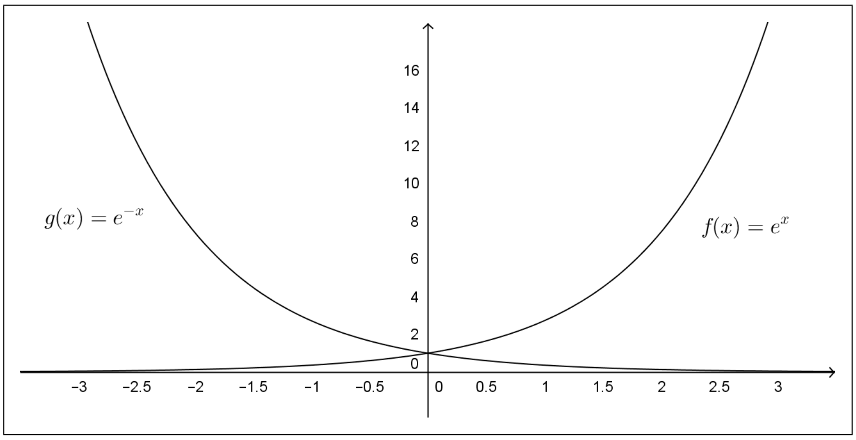


Figura 2.4 – Gráfico das funções $f(x) = e^x$ e $g(x) = e^{-x}$

2.2.1 Função densidade

Para que uma função $f(x)$ seja uma função densidade, sua integral no domínio de definição tem que ser igual a 1. Considerando, então, a função $f(x) = e^{-\lambda x}$ para $x > 0$, temos que

$$\int_0^\infty e^{-\lambda x} dx = \left(-\frac{1}{\lambda} e^{-\lambda x} \right)_0^\infty = \frac{1}{\lambda}$$

Logo,

$$\lambda \int_0^\infty e^{-\lambda x} dx = 1 \Rightarrow \int_0^\infty \lambda e^{-\lambda x} dx = 1$$

e, portanto, $f(x) = \lambda e^{-\lambda x}$ define uma função de densidade de probabilidade para $x > 0$.

DEFINIÇÃO **Densidade exponencial**

Diz-se que uma variável aleatória contínua X tem distribuição exponencial com parâmetro λ se sua função densidade de probabilidade é dada por

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

Como essa função depende apenas do valor de λ, esse é o parâmetro da densidade exponencial.

Usaremos a seguinte notação para indicar que uma variável aleatória tem distribuição exponencial com parâmetro λ: $X \sim \exp(\lambda)$. Na Figura 2.5 temos o gráfico da densidade exponencial para $\lambda = 2$.

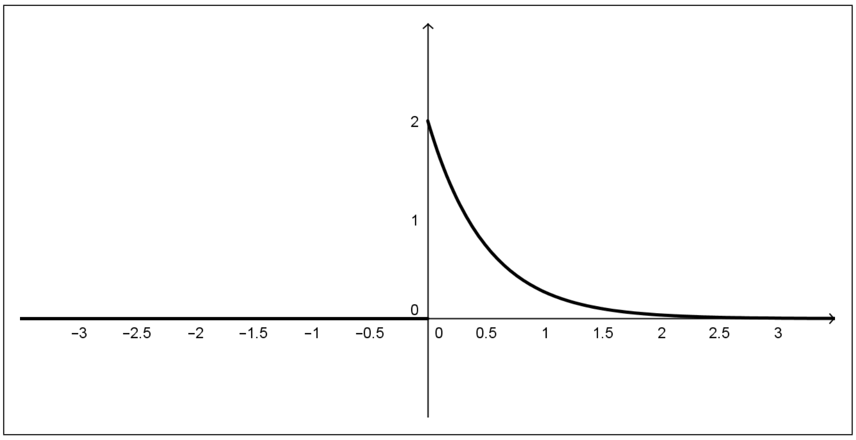


Figura 2.5 – Densidade exponencial com $\lambda = 2$

2.2.2 Função de distribuição

Por definição, temos que

$$F(x) = \Pr(X \leq x) = \int_0^x f(t)\,dt = \int_0^x \lambda e^{-\lambda t} dt = -e^{-\lambda t}\Big|_0^x = -\left(e^{-\lambda x} - 1\right)$$

ou seja

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{se } x > 0 \\ 0 & \text{se } x \leq 0 \end{cases} \tag{2.4}$$

Na Figura 2.6 temos o gráfico da função de distribuição para $\lambda = 2$.

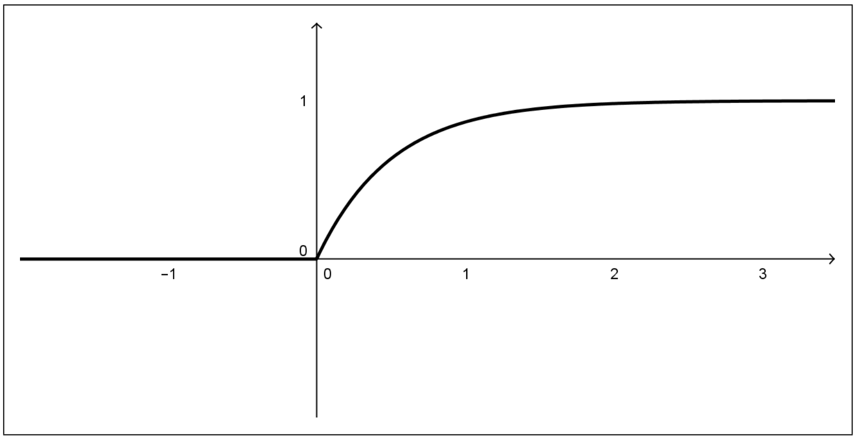


Figura 2.6 – Função de distribuição da densidade exponencial com $\lambda = 2$

2.2.3 Alguns resultados sobre a função exponencial

No cálculo dos momentos da densidade exponencial serão necessários alguns resultados sobre a função exponencial que apresentaremos a seguir.

O resultado crucial é

$$\lim_{x\to\infty} x^k e^{-x} = 0 \tag{2.5}$$

Vamos demonstrar esse resultado por indução usando a regra de L'Hôpital. Consideremos inicialmente o caso em que $k = 1$. Então

$$\lim_{x\to\infty} xe^{-x} = \lim_{x\to\infty} \frac{x}{e^x}$$

que tem a forma $\frac{\infty}{\infty}$ e, portanto, podemos aplicar a regra de L'Hôpital, que diz que

$$\lim_{x\to\infty} xe^{-x} = \lim_{x\to\infty} \frac{x}{e^x} = \lim_{x\to\infty} \frac{x'}{(e^x)'} = \lim_{x\to\infty} \frac{1}{e^x} = 0$$

Logo, o resultado vale para $k = 1$. Suponhamos verdadeiro para qualquer $k > 1$; vamos mostrar que vale para $k + 1$. De fato:

$$\lim_{x\to\infty} x^{k+1}e^{-x} = \lim_{x\to\infty} \frac{x^{k+1}}{e^x} = \lim_{x\to\infty} \frac{\left(x^{k+1}\right)'}{(e^x)'} = \lim_{x\to\infty} \frac{(k+1)\,x^k}{e^x} = (k+1) \lim_{x\to\infty} \frac{x^k}{e^x} = (k+1)\times 0 = 0$$

pela hipótese de indução. De maneira análoga, prova-se um resultado mais geral dado por:

$$\lim_{x\to\infty} x^k e^{-\lambda x} = 0 \qquad \forall k > 0 \text{ e } \lambda > 0 \tag{2.6}$$

A interpretação desse resultado pode ser vista na Figura 2.7 para o caso de x^3: para valores de x maiores que a abcissa do ponto de interseção representado em vermelho, o valor de e^x é sempre maior que x^3 e, portanto o limite de $\frac{x^3}{e^x}$ é zero. Dito de outra forma, a função exponencial cresce muito mais rapidamente que qualquer função polinomial.

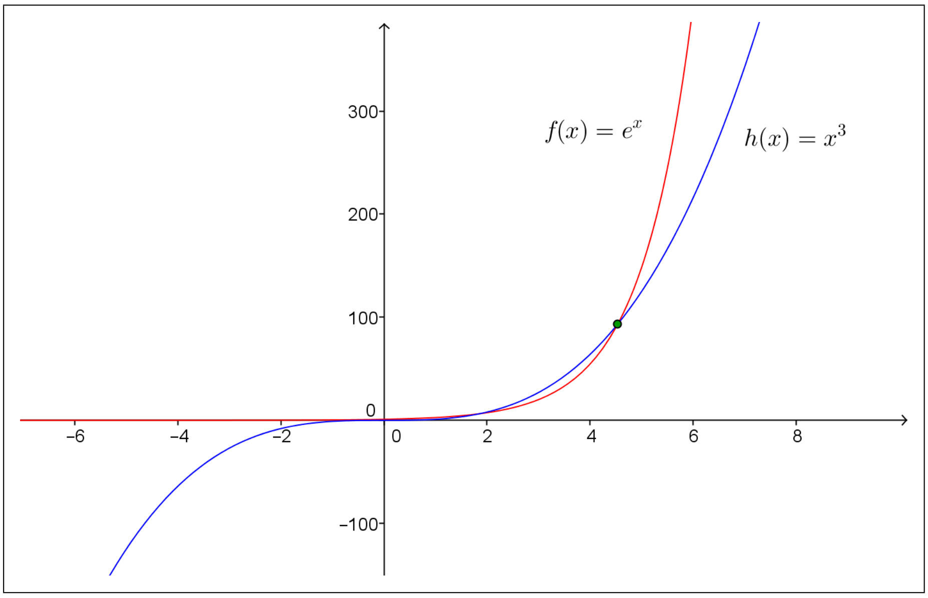


Figura 2.7 – Ilustração do resultado $\lim_{x\to\infty} \dfrac{x^3}{e^x} = 0$

2.2.4 Esperança e variância

O cálculo dos momentos da distribuição exponencial se faz com auxílio de integração por partes. A esperança é:

$$\mathrm{E}(X) = \int_0^\infty x\lambda e^{-\lambda x} dx$$

Definindo

- $u = x \Rightarrow du = dx$;
- $dv = \lambda e^{-\lambda x} dx \Rightarrow v = -e^{-\lambda x}$

  O método de integração por partes nos dá que:

$$-xe^{-\lambda x}\Big|_0^\infty = \int_0^\infty x\lambda e^{-\lambda x} dx + \int_0^\infty \left(-e^{-\lambda x}\right) dx$$

Pelo resultado (2.5), o lado esquerdo desta última igualdade é zero. Logo,

$$0 = \mathrm{E}(X) + \frac{1}{\lambda} e^{-\lambda x}\Big|_0^\infty \Rightarrow 0 = \mathrm{E}(X) + \left(0 - \frac{1}{\lambda}\right)$$

ou seja,

$$\mathrm{E}\,(X) = \frac{1}{\lambda} \tag{2.7}$$

Desse resultado segue que

$$\int_0^\infty x\lambda e^{-\lambda x} dx = \frac{1}{\lambda} \Rightarrow \int_0^\infty xe^{-\lambda x} dx = \frac{1}{\lambda^2} \tag{2.8}$$

Vamos calcular o segundo momento de uma variável aleatória exponencial.

$$\mathrm{E}(X^2) = \int_0^\infty x^2 \lambda e^{-\lambda x} dx$$

Seguindo raciocínio análogo ao empregado no cálculo da esperança, vamos definir:

- $u = x^2 \Rightarrow du = 2xdx$;
- $dv = \lambda e^{-\lambda x} dx \Rightarrow v = -e^{-\lambda x}$

Logo,

$$-x^2 e^{-\lambda x}\Big|_0^\infty = \int_0^\infty x^2 \lambda e^{-\lambda x} dx + \int_0^\infty \left(-2xe^{-\lambda x}\right) dx \Rightarrow 0 = E\left(X^2\right) - 2\int_0^\infty xe^{-\lambda x} dx$$

Por (2.8), resulta que

$$\mathrm{E}\left(X^2\right) = \frac{2}{\lambda^2} \tag{2.9}$$

e, portanto:

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} \Rightarrow \mathrm{Var}\,(X) = \frac{1}{\lambda^2} \tag{2.10}$$

Resumindo:

$$X \sim \exp(\lambda) \Longrightarrow \begin{cases} \mathrm{E}(X) = \dfrac{1}{\lambda} \\ \\ \mathrm{Var}(X) = \dfrac{1}{\lambda^2} \end{cases} \tag{2.11}$$

2.2.5 Parametrização alternativa

É possível parametrizar a densidade exponencial em termos de um parâmetro $\beta = \frac{1}{\lambda}$. Neste caso,

$$\begin{aligned} f(x) &= \frac{1}{\beta} e^{-x/\beta} \qquad x > 0; \beta > 0 \\ \mathrm{E}(X) &= \beta \\ \mathrm{E}(X^2) &= 2\beta^2 \\ \mathrm{Var}(X) &= \beta^2 \end{aligned}$$

Essa parametrização alternativa é mais interessante, uma vez que o valor médio é igual ao parâmetro, e será utilizada deste ponto em diante.

2.2.6 Exemplos

EXEMPLO 2.3

Seja X uma variável aleatória exponencial com média 4. Calcule

(a) $\mathrm{P}(X > 1)$

(b) $\mathrm{Pr}(1 \leq X \leq 2)$

Solução

(a) A função densidade de X é $f_X(x) = \frac{1}{4} e^{-x/4}$ e a função de distribuição é $F(x) = 1 - e^{-x/4}$. Podemos resolver essa questão por integração da função densidade

$$\mathrm{P}(X > 1) = \int_1^\infty f(x)dx = \int_1^\infty \frac{1}{4} e^{-\frac{x}{4}} dx = -e^{-\frac{x}{4}}\Big|_1^\infty = 0 - (-e^{0,25}) = e^{-0,25} = 0,7788$$

ou pelo uso direto da função de distribuição

$$\mathrm{P}(X > 1) = 1 - \mathrm{Pr}(X \leq 1) = 1 - F(1) = 1 - [1 - e^{-1/4}] = e^{-0,25} = 0,7788$$

(b)

$$\begin{aligned} \mathrm{P}(1 \leq X \leq 2) &= \mathrm{P}(X \leq 2) - \mathrm{Pr}(X < 1) \\ &= \mathrm{P}(X \leq 2) - \mathrm{P}(X \leq 1) \\ &= F(2) - F(1) = [1 - e^{-2/4}] - [1 - e^{-1/4}] \\ &= e^{-0,25} - e^{-0,5} = 0,17227 \end{aligned}$$

EXEMPLO 2.4

Seja $X \sim \exp(\beta)$. Calcule $\mathrm{P}(X > E(X))$.

Solução

$$\mathrm{P}(X > E(X)) = 1 - \mathrm{P}(X \leq E(X)) = 1 - F(\mathrm{E}(X)) = 1 - \left[1 - e^{-\beta/\beta}\right] = e^{-1}$$

Note que essa é a probabilidade de uma variável aleatória exponencial ser maior que o seu valor médio; o que mostramos é que essa probabilidade é constante, qualquer que seja o parâmetro.

2.3 Distribuição gama

A distribuição exponencialé um caso particular da distribuição gama, que utiliza a *função gama*, cuja definição apresentamos a seguir.

2.3.1 A função gama

A função gama é definida pela seguinte integral:

$$\Gamma(\alpha) = \int_0^\infty e^{-x} x^{\alpha-1} dx \qquad \alpha \geq 1$$

Note que o argumento da função é α, que aparece no expoente da variável de integração x.

A função gama tem a seguinte propriedade recursiva: $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$. Para demonstrar esse resultado, usaremos integração por partes.

$$\Gamma(\alpha + 1) = \int_0^\infty e^{-x} x^{\alpha} dx$$

Fazendo

- $u = x^\alpha \Rightarrow du = \alpha x^{\alpha-1}$
- $dv = e^{-x}dx \Rightarrow v = -e^{-x}$

resulta

$$\begin{aligned}
\Gamma(\alpha + 1) &= -x^\alpha e^{-x}\Big|_0^\infty - \int_0^\infty -e^{-x}\alpha x^{\alpha-1} dx \Rightarrow \\
\Gamma(\alpha + 1) &= 0 + \alpha \int_0^\infty e^{-x} x^{\alpha-1} dx \Rightarrow \\
\Gamma(\alpha + 1) &= \alpha\Gamma(\alpha)
\end{aligned}$$

Aqui usamos o resultado dado em (2.5).

Vamos trabalhar, agora, com $\alpha = n$ inteiro.

$$\Gamma(1) = \int_0^\infty e^{-x} x^{1-1} dx = \int_0^\infty e^{-x} dx = 1$$

$$\begin{aligned}
\Gamma(2) &= 1 \times \Gamma(1) = 1 = 1! \\
\Gamma(3) &= 2 \times \Gamma(2) = 2 \times 1 = 2! \\
\Gamma(4) &= 3 \times \Gamma(3) = 3 \times 2 \times 1 = 3! \\
\Gamma(5) &= 4 \times \Gamma(4) = 4 \times 3 \times 2 \times 1 = 4!
\end{aligned}$$

Em geral, se n é inteiro,

$$\Gamma(n) = (n-1)! \tag{2.12}$$

2.3.2 A distribuição gama

DEFINIÇÃO Densidade gama

Diz-se que uma variável aleatória tem distribuição gama com parâmetros α e β se sua função densidade de probabilidade é dada por

$$f(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} & \text{se } x > 0 \\ 0 & \text{se } x \leq 0 \end{cases} \tag{2.13}$$

Note que, quando $\alpha = 1$, resulta a densidade exponencial com parâmetro β, ou seja, a densidade exponencial é um caso particular da densidade gama. Note que estamos usando a parametrização alternativa da densidade exponencial.

Para verificar que a função dada em (2.13) realmente define uma função densidade, notamos inicialmente que $f(x) \geq 0$. Além disso,

$$\int_0^\infty f(x)dx = \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{-x/\beta} dx$$

Fazendo a mudança de variável $\dfrac{x}{\beta} = t$ resulta

$$\begin{aligned} x &= \beta t \\ dx &= \beta dt \\ x &= 0 \Rightarrow t = 0 \\ x &= \infty \Rightarrow t = \infty \end{aligned}$$

e, portanto,

$$\begin{aligned} \int_0^\infty f(x)dx &= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{-x/\beta} dx = \\ &= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty (\beta t)^{\alpha-1} e^{-t} \beta dt \\ &= \frac{1}{\Gamma(\alpha)\beta^\alpha} \beta^\alpha \int_0^\infty t^{\alpha-1} e^{-t} dt \\ &= \frac{1}{\Gamma(\alpha)} \Gamma(\alpha) = 1 \end{aligned}$$

Logo, as duas condições para uma função densidade são satisfeitas. Usaremos a notação $X \sim gama(\alpha; \beta)$ para indicar que a variável aleatória X tem distribuição gama com parâmetros α, β.

2.3.3 O gráfico da distribuição gama

Qualquer que seja o valor do parâmetro β, temos que

$$\lim_{x \to \infty} f(x) = 0$$

$$\lim_{x \to 0} f(x) = \begin{cases} 0 & \text{se } \alpha \geq 1 \\ \infty & \text{se } \alpha < 1 \end{cases}$$

Em cada uma das Figuras 2.8 e 2.9, o parâmetro β está fixo e temos o gráfico da função densidade para $\alpha = 1, 2, 4$. Nas Figuras 2.10 e 2.11, fixamos o parâmetro α e variamos β. Analisando esses gráficos,podemos ver que α tem grande influência sobre a forma da distribuição, enquanto β afeta mais fortemente a dispersão. Por essas razões, α é dito *parâmetro de forma* da densidade gama e β é o *parâmetro de escala*.

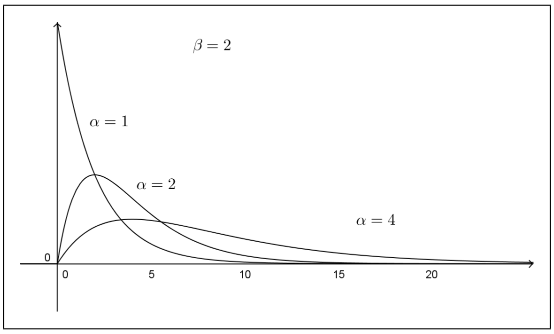


Figura 2.8 – Efeito do parâmetro α – $\beta = 2; \alpha = 1, 2, 4$

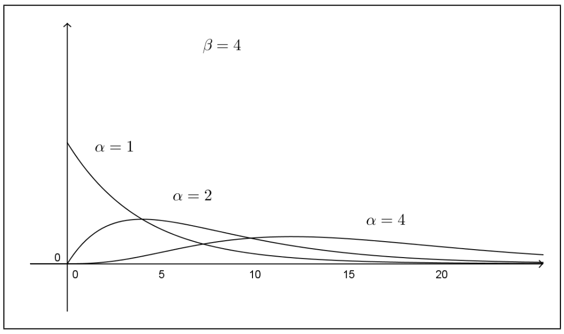


Figura 2.9 – Efeito do parâmetro α – $\beta = 4; \alpha = 1, 2, 4$

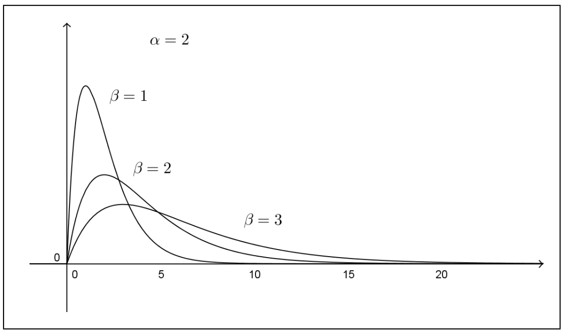


Figura 2.10 – Efeito do parâmetro β – $\alpha = 2; \beta = 1, 2, 3$

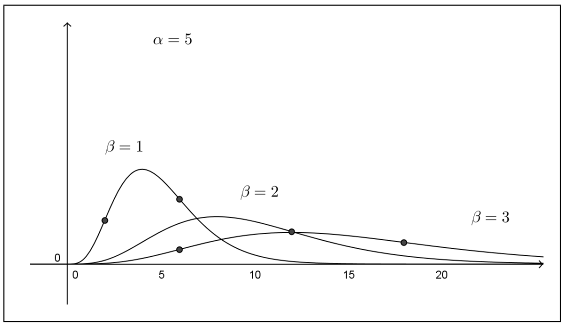


Figura 2.11 – Efeito do parâmetro β – $\alpha = 5; \beta = 1, 2, 3$

Uma análise das derivadas primeira e segunda da função densidade gama permite-nos chegar às seguintes conlusões sobre a forma do gráfico:

1. $\alpha \leq 1$
   (a) estritamente decrescente
   (b) côncava para cima
2. $\alpha > 1$
   (a) crescente se $x < \beta(\alpha - 1)$

(b) decrescente se $x > \beta(\alpha - 1)$

(c) máximo em $x = \beta(\alpha - 1)$

(d) $\alpha \leq 2$

i. único ponto de inflexão em $x = \beta\sqrt{\alpha - 1}(\sqrt{\alpha - 1} + 1)$
ii. côncava para baixo se $x < \beta\sqrt{\alpha - 1}(\sqrt{\alpha - 1} + 1)$
iii. côncava para cima se $x > \beta\sqrt{\alpha - 1}(\sqrt{\alpha - 1} + 1)$

(e) $\alpha > 2$

i. dois pontos de inflexão: $x = \beta\sqrt{\alpha - 1}(\sqrt{\alpha - 1} - 1)$ e $x = \beta\sqrt{\alpha - 1}(\sqrt{\alpha - 1} + 1)$
ii. côncava para cima se $x < \beta\sqrt{\alpha - 1}(\sqrt{\alpha - 1} - 1)$
iii. côncava para baixo se $\beta\sqrt{\alpha - 1}(\sqrt{\alpha - 1} - 1) < x < \beta\sqrt{\alpha - 1}(\sqrt{\alpha - 1} + 1)$
iv. côncava para cima se $x > \beta\sqrt{\alpha - 1}(\sqrt{\alpha - 1} + 1)$

2.3.4 Esperança e variância

Se $X \sim gama(\alpha; \beta)$, então

$$\begin{aligned}
\mathrm{E}(X) &= \int_0^\infty x f(x) dx = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x x^{\alpha-1} e^{-x/\beta} dx \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^\alpha e^{-x/\beta} dx
\end{aligned}$$

Fazendo a mesma mudança de variável já usada anteriormente $\frac{x}{\beta} = t$ temos que

$$\begin{aligned}
\mathrm{E}(X) &= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty (\beta t)^\alpha e^{-t} \beta dt \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \beta^{\alpha+1} \int_0^\infty t^\alpha e^{-t} dt \\
&= \frac{\beta}{\Gamma(\alpha)} \int_0^\infty t^\alpha e^{-t} dt \\
&= \frac{\beta}{\Gamma(\alpha)} \Gamma(\alpha + 1) \\
&= \frac{\beta}{\Gamma(\alpha)} \alpha \Gamma(\alpha)
\end{aligned}$$

ou seja,

$$X \sim gama(\alpha, \beta) \Rightarrow \mathrm{E}(X) = \alpha\beta$$

De modo análogo, vamos calcular o segundo momento da densidade gama.

$$\begin{aligned}
\mathrm{E}(X^2) &= \int_0^\infty x^2 f(x) dx = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^2 x^{\alpha-1} e^{-x/\beta} dx \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha+1} e^{-x/\beta} dx
\end{aligned}$$

Fazendo a mesma mudança de variável usada anteriormente $\frac{x}{\beta} = t$ temos que

$$\begin{aligned}
\mathrm{E}(X^2) &= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty (\beta t)^{\alpha+1} e^{-t} \beta dt \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \beta^{\alpha+2} \int_0^\infty t^{\alpha+1} e^{-t} dt \\
&= \frac{\beta^2}{\Gamma(\alpha)} \int_0^\infty t^{\alpha+1} e^{-t} dt \\
&= \frac{\beta^2}{\Gamma(\alpha)} \Gamma(\alpha+2) \\
&= \frac{\beta^2}{\Gamma(\alpha)} (\alpha+1)\Gamma(\alpha+1) \\
&= \frac{\beta^2}{\Gamma(\alpha)} (\alpha+1)\alpha\Gamma(\alpha) \\
&= \beta^2(\alpha+1)\alpha
\end{aligned}$$

Logo,

$$\mathrm{Var}(X) = \beta^2(\alpha+1)\alpha - (\alpha\beta)^2 = \alpha^2\beta^2 + \alpha\beta^2 - \alpha^2\beta^2 = \alpha\beta^2$$

Resumindo:

$$X \sim gama(\alpha, \beta) \Longrightarrow \begin{cases} \mathrm{E}(X) = \alpha\beta \\ \\ \mathrm{Var}(X) = \alpha\beta^2 \end{cases} \tag{2.14}$$

2.3.5 Função de distribuição acumulada

A função de distribuição da gama envolve a função gama incompleta e não será objeto de estudo neste curso.

2.4 A distribuição Erlang

Quando o parâmetro de forma α da densidade gama é um inteiro positivo, a distribuição gama é conhecida como *distribuição de Erlang* e sua função densidade é

$$f(x) = \begin{cases} \dfrac{x^{k-1} e^{-x/\beta}}{(k+1)!\beta^k} & \text{se } x > 0 \\ \\ 0 & \text{se } x \leq 0 \end{cases} \tag{2.15}$$

2.5 A distribuição qui-quadrado

Quando, na distribuição gama, o parâmetro de forma é igual a $\frac{n}{2}$, com n inteiro positivo, e o parâmetro de escala é $\beta = 2$ resulta a distribuição qui-quadrado com n graus de liberdade, cuja densidade é

$$f(x) = \frac{1}{\Gamma\left(\frac{n}{2}\right) 2^{n/2}} x^{n/2-1} e^{-x/2} \qquad \text{se } x > 0 \tag{2.16}$$

Usaremos a seguinte notação para indicar que X tem distribuição qui-quadrado com n graus de liberdade: $X \sim \chi^2_n$. Usando os resultados dados em (2.14), temos

$$X \sim \chi^2_n \Longrightarrow \begin{cases} E(X) = \frac{n}{2} \cdot 2 = n \\ Var(X) = \frac{n}{2} \cdot 2^2 = 2n \end{cases}$$

Na Figura 2.12 apresenta-se o gráfico da densidade qui-quadrado com 3 graus de liberdade.

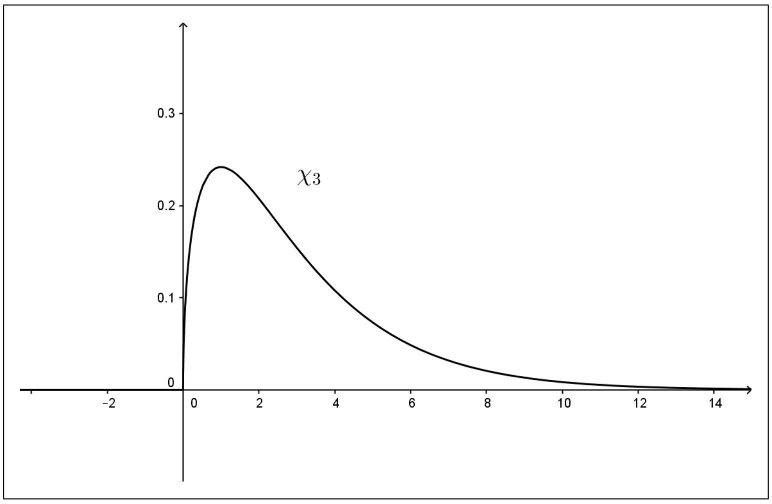


Figura 2.12 – Densidade qui-quadrado com 3 g.l.

2.5.1 A tabela da qui-quadrado

Como o parâmetro da distribuição qui-quadrado é o número de graus de liberdade n, seria necessária uma tabela para cada valor de n, uma vez que não existe qualquer relação entre diferentes distribuições qui-quadrado, como ocorre no caso da distribução normal. Os programas computacionais de estatística calculam probabilidades associadas a qualquer distribuição χ^2. Mas nos livros didáticos, é comum apresentar uma tabela da distribuição χ^2_n que envolve os *valores críticos*, ou seja, valores que deixam determinada probabilidade acima deles. Mais precisamente, o valor crítico da χ^2_n associado à probabilidade α é o valor $\chi^2_{n;\alpha}$ tal que

$$\mathrm{P}(\chi^2_n > \chi^2_{n;\alpha}) = \alpha$$

Veja a Figura 2.13.

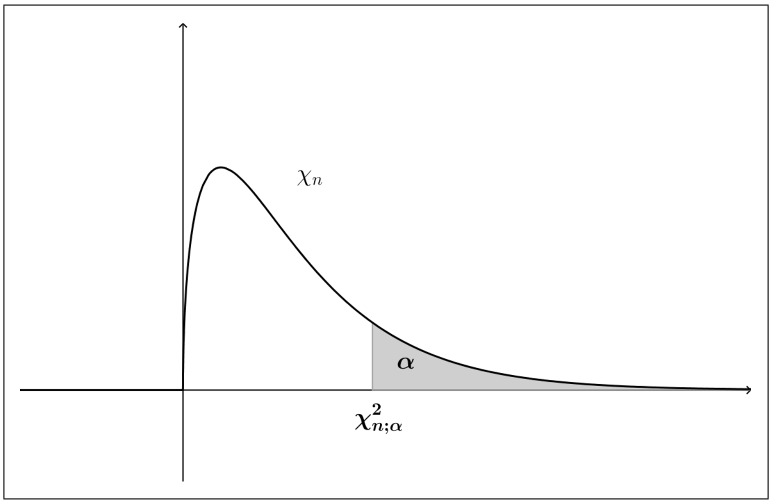


Figura 2.13 – Valor crítico $\chi_{n,\alpha}$ da distribuição χ^2_n: $P(\chi^2_n > \chi^2_{n,\alpha}) = \alpha$

No Apêndice **??**, apresentamos a Tabela 3, que é uma apresentação usual dos valores críticos da distribuição χ^2_n. Nesta tabela, cada linha corresponde a um número diferente de graus de liberdade e cada coluna corresponde a uma área α na cauda superior. No corpo da tabela temos o valor crítico $\chi^2_{n;\alpha}$.

EXEMPLO 2.5 Utilizando a Tabela 3

Consulte a Tabela 3 da distribuição qui-quadrado para resolver os seguintes exercícios.

(a) Determine o valor de k tal que $P(\chi^2_8 \geq k) = 0,05$.

(b) Determine o valor de k tal que $P(\chi^2_{12} \leq k) = 0,10$.

(c) Se $p = P(\chi^2_{10} \geq 18,123)$, determine valores p_1 e p_2 tais que $p_1 < p < p_2$.

(d) Determine valores k_1 e k_2 tais que $k_1 < k < k_2$ quando k é tal que $P(\chi^2_6 \geq k) = 0,036$.

Solução

O primeiro ponto a observar é que a tabela trabalha com probabilidades na cauda superior da distribuição.

(a) Temos que olhar na linha correspondente a 8 graus de liberdade e coluna correspondente à probabilidade 0,05: $k = \chi^2_{8;0,05} = 15,507$.

(b) O problema pede probabilidade 0,10 na cauda inferior, o que corresponde à probabilidade 0,90 na causa superior, ou seja, $k = \chi^2_{12;0,90} = 7,042$

(c) Olhando na linha correspondente a 10 g.l., podemos ver que a abcissa 18,123 está entre 15,987 e 18,307. As probabilidades acima dessas duas abcissas são 0,10 e 0,05, respectivamente. Veja as Figuras 2.14 e 2.15. Concluímos, então, que a probabilidade acima de 18,123 é um valor p tal que $0,05 < P < 0,10$ - o comando do excel "`=1-DIST.QUIQUA(18,123;10;1)`" retorna a probabilidade exata 0,052923914.

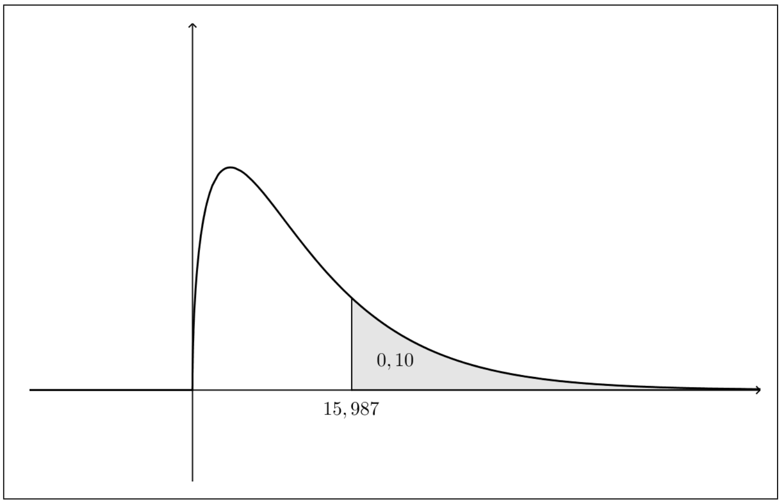


Figura 2.14 – $P(\chi^2 10 > 15,987) = 0,10$

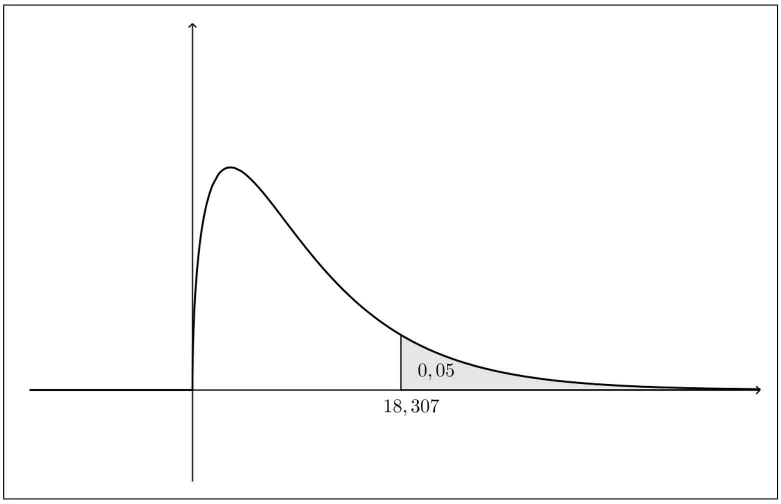


Figura 2.15 – $P(\chi^2 10 > 18,307) = 0,05$

(d) Olhando na na tabela, podemos ver que a probabilidade 0,036 está entre as probabilidades 0,025 e 0,05, ou seja, $0,025 < 0,036 < 0,05$. Olhando na linha correspondente a 6 g.l., podemos ver que a abcissa correspondente à probabilidade 0,025 é 14,449 e a abcissa correspondente à probabilidade 0,05 é 12,592. Veja as Figuras 2.16 a 2.18. Concluímos, então, que a probabilidade acima de $12,592 < k < 14,449$. 18,123 é um valor p tal que $0,05 < P < 0,10$ - o comando do excel "`=INV.QUIQUA(1-0,036;6)`" retorna a abcissa exata 13,48119895.

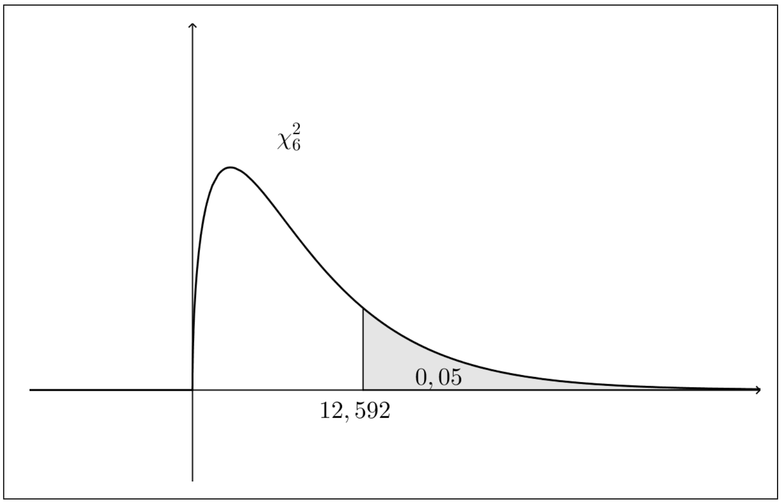


Figura 2.16 – $P(\chi^2_6 > 12,592) = 0,05$

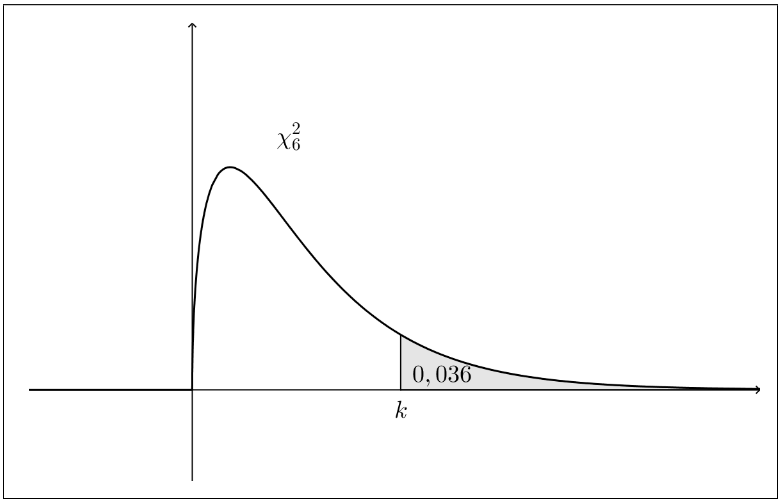


Figura 2.17 – $P(\chi^26 > k) = 0,036$

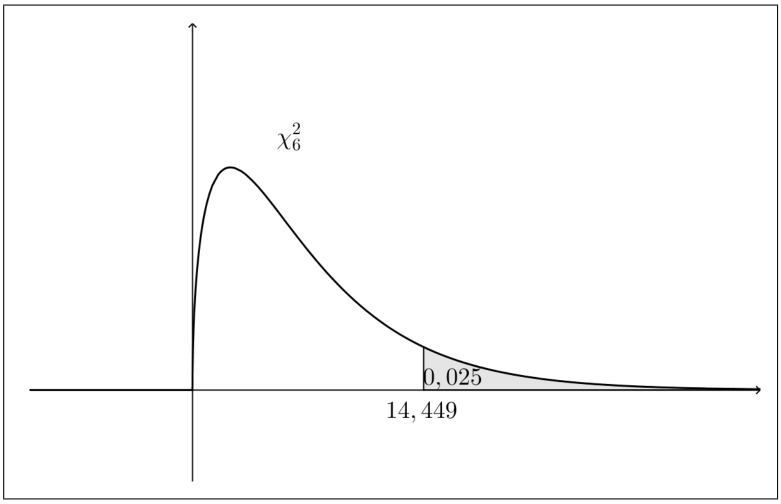


Figura 2.18 – $P(\chi^26 > 14,449) = 0,025$

2.6 Distribuição de Pareto

2.6.1 Função densidade

DEFINIÇÃO Densidade de Pareto

Uma variável aleatória X tem distribuição de Pareto com parâmetros $\alpha > 0$ e $b > 0$ se sua função densidade de probabilidade é dada por

$$f(x) = \begin{cases} \dfrac{\alpha}{b}\left(\dfrac{b}{x}\right)^{\alpha+1} & \text{se } x \geq b \\ 0 & \text{se } x < b \end{cases}$$

Para mostrar que $f(x)$ realmente define uma função densidade de probabilidade resta provar que a integral é 1, uma vez que $f(x) \geq 0$.

$$\int_b^\infty \frac{\alpha}{b}\left(\frac{b}{x}\right)^{\alpha+1} dx = \alpha b^\alpha \int_b^\infty x^{-\alpha-1} dx = \alpha b^\alpha \left.\frac{x^{-\alpha}}{-\alpha}\right|_b^\infty$$

Essa integral converge apenas se $-\alpha < 0$ ou equivalentemente, $\alpha > 0$, pois nesse caso $\lim_{x\to\infty} x^{-\alpha} = \lim_{x\to\infty} \frac{1}{x^\alpha} = 0$. Satisfeita esta condição, temos que

$$\alpha b^\alpha \left.\frac{x^{-\alpha}}{-\alpha}\right|_b^\infty = 0 - \alpha b^\alpha \frac{b^{-\alpha}}{-\alpha} = 1$$

Na **Figura 2.19** ilustra-se a densidade de Pareto para $\alpha = 3$ e $b = 2$.

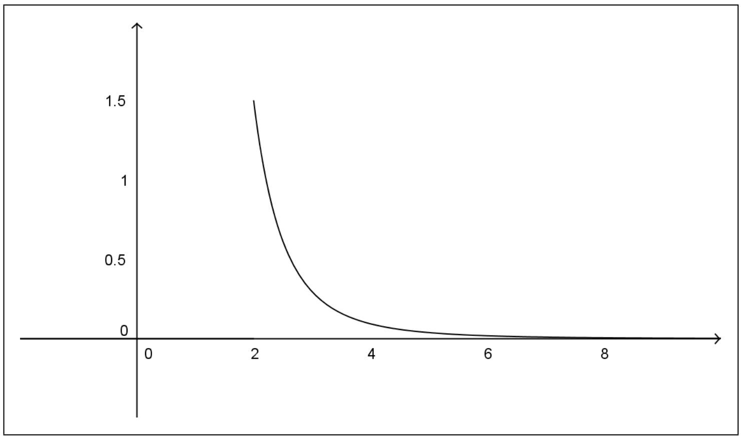


Figura 2.19 – Densidade de Pareto com parâmetros $\alpha = 3$ e $b = 2$

2.6.2 Esperança e variância

Se $X \sim Pareto(\alpha, b)$ então

$$\mathrm{E}(X) = \int_b^\infty x\frac{\alpha}{b}\left(\frac{b}{x}\right)^{\alpha+1} dx = \alpha b^\alpha \int_b^\infty x^{-\alpha} dx = \alpha b^\alpha \left.\frac{x^{-\alpha+1}}{-\alpha+1}\right|_b^\infty$$

Para que essa integral convirja, temos que ter $-\alpha+1 < 0$, ou $\alpha > 1$. Satisfeita esta condição,

$$\mathrm{E}(X) = \alpha b^{\alpha}\left(0 - \frac{b^{-\alpha+1}}{-\alpha+1}\right) = \frac{-\alpha b^{\alpha-\alpha+1}}{1-\alpha} = \frac{\alpha b}{\alpha - 1}$$

O segundo momento é

$$\mathrm{E}(X^2) = \int_b^{\infty} x^2 \frac{\alpha}{b}\left(\frac{b}{x}\right)^{\alpha+1} dx = \alpha b^{\alpha}\int_b^{\infty} x^{-\alpha+1} dx = \alpha b^{\alpha} \left.\frac{x^{-\alpha+2}}{-\alpha+2}\right|_b^{\infty}$$

Para que essa integral convirja, temos que ter $-\alpha+2 < 0$, ou $\alpha > 2$. Satisfeita esta condição,

$$\mathrm{E}(X) = \alpha b^{\alpha}\left(0 - \frac{b^{-\alpha+2}}{-\alpha+2}\right) = \frac{-\alpha b^{\alpha-\alpha+2}}{2-\alpha} = \frac{\alpha b^2}{\alpha - 2}$$

Logo,

$$\begin{aligned}
\mathrm{Var}(X) &= \frac{\alpha b^2}{\alpha-2} - \left(\frac{\alpha b}{\alpha-1}\right)^2 = \frac{\alpha b^2 (\alpha-1)^2 - \alpha^2 b^2 (\alpha-2)}{(\alpha-1)^2(\alpha-2)} \\
&= \frac{\alpha b^2\left[\alpha^2 - 2\alpha + 1 - \alpha(\alpha-2)\right]}{(\alpha-1)^2(\alpha-2)} = \frac{\alpha b^2\left[\alpha^2 - 2\alpha + 1 - \alpha^2 + 2\alpha\right]}{(\alpha-1)^2(\alpha-2)} \\
&= \frac{\alpha b^2}{(\alpha-1)^2(\alpha-2)}
\end{aligned}$$

Resumindo:

$$X \sim Pareto(\alpha, b) \Longrightarrow \begin{cases} \mathrm{E}(X) = \dfrac{\alpha b}{\alpha-1} & \text{se } \alpha > 1 \\ \mathrm{Var}(X) = \dfrac{\alpha b^2}{(\alpha-1)^2(\alpha-2)} & \text{se } \alpha > 2 \end{cases} \tag{2.17}$$

2.6.3 Função de distribuição

Por definição, $F(x) = \mathrm{P}(X \leq x) = 0$ se $x < b$. Para $x \geq b$,

$$\begin{aligned}
F(x) &= \Pr(X \leq x) = \int_b^x \frac{\alpha}{b}\left(\frac{b}{t}\right)^{\alpha+1} dt = \alpha b^{\alpha}\int_b^x t^{-\alpha-1} dx = \alpha b^{\alpha} \left.\frac{t^{-\alpha}}{-\alpha}\right|_b^x \\
&= -b^{\alpha}\left(x^{-\alpha} - b^{-\alpha}\right) = 1 - \left(\frac{b}{x}\right)^{\alpha}
\end{aligned}$$

$$X \sim Pareto(\alpha, b) \Longrightarrow F(x) = \begin{cases} 0 & \text{se } x < b \\ 1 - \left(\frac{b}{x}\right)^{\alpha} & \text{se } x \geq b \end{cases} \tag{2.18}$$

Na Figura 2.20 ilustra-se a função de distribuição de Pareto com parâetros $\alpha = 3$ e $b = 2$.

2.7 Exercícios propostos

2.1 Você está interessado em dar um lance em um leilão de um lote de terra. Você sabe que existe um outro licitante. Pelas regras estabelecidas para este leilão, o lance mais alto acima de R$ 100.000,00 será aceito. Suponha que o lance do seu competidor seja uma variável aleatória uniformemente distribuída entre R$ 100.000,00 e R$ 150.000,00.

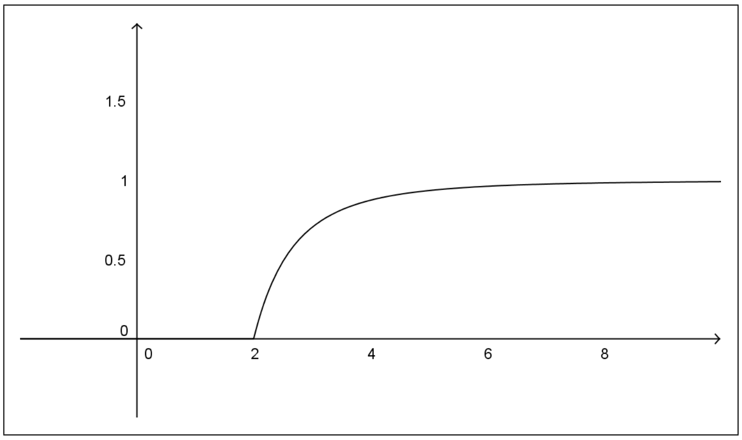


Figura 2.20 – Função de distribuição de Pareto com parâmetros $\alpha = 3$ e $b = 2$

(a) Se você der um lance de R$120.000,00, qual é a probabilidade de você ficar com o lote? (Resp.: $0,4$)

(b) Se você der um lance de R$140.000,00, qual é a probabilidade de você ficar com o lote? (Resp.: $0,8$)

2.2 Seja X uma variável aleatória com distribuição exponencial de média 8. Calcule as seguintes probabilidades:

(a) $P(X > 10)$ (Resp.: $0,286505$)

(b) $P(X > 8)$ (Resp.: 0.36788)

(c) $P(5 < X < 11)$ (Resp.: $0,28242$)

2.3 O tempo entre chegadas de automóveis num lava-jato é distribuído exponencialmente, com uma média de 12 minutos.

(a) Qual é a probabilidade de que o tempo entre chegadas de veículos neste lava-jato seja maior que 10 minutos? (Resp.: $0,43460$)

(b) Qual é a probabilidade de que o tempo entre chegadas de veículos neste lava-jato seja menor que 8 minutos? (Resp.: $0,48658$)

2.4 Consulte a Tabela 3 (Apêndice **??**) da distribuição qui-quadrado para resolver os seguintes exercícios.

(a) Determine o valor de k tal que $P(\chi^2_{19} \geq k) = 0,025$. (Resp.: $32,852$)

(b) Determine o valor de k tal que $P(\chi^2_{17} \leq k) = 0,05$. (Resp.: $8,672$)

(c) Se $p = P(\chi^2_{15} \geq 13,43)$, determine valores p_1 e p_2 tais que $p_1 < p < p_2$. (Resp.: $p_1 = 0,50; p_2 = 0,70$; valor exato de p é 0,569122081.)

(d) Determine valores k_1 e k_2 tais que $k_1 < k < k_2$ quando k é tal que $P(\chi^2_{26} \geq k) = 0,017$. (Resp.: $k_1 = 42,856; k_2 = 45,642$. O valor exato de k é 43,52333776.)

Capítulo 3

Funções de Variáveis Aleatórias Contínuas

Dada uma variável aleatória contínua X com função de densidade $f_X(x)$, muitas vezes estamos interessados em conhecer a densidade de uma outra variável aleatória $Y = g(x)$ definida como uma função de X. Vamos apresentar alguns exemplos que ilustram dois métodos comuns. O primeiro utiliza a relação entre a função densidade e a função de distribuição e o segundo se aplica quando a função g é inversível.

3.1 Método da função de distribuição

EXEMPLO 3.1 Transformações da uniforme

Seja $X \sim Unif(-1, 1)$. Vamos calcular a função densidade das novas variáveis aleatórias $Y = g(X) = X^2$ e $W = h(X) = |X|$. Para isso usaremos a função de distribuição e sua relação com a função de densidade dada por

$$f_X(x) = F'_X(x) \tag{3.1}$$

Temos que

$$f_X(x) = \begin{cases} \frac{1}{2} & -1 < x < 1 \\ 0 & x \leq -1 \text{ ou } x \geq 1 \end{cases}$$

$$-1 < x < 1 \Rightarrow \begin{cases} 0 \leq g(x) < 1 \\ 0 \leq h(x) < 1 \end{cases}$$

Para calcular a função de densidade de probabilidade de $Y = g(X) = X^2$ devemos notar que

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(X^2 \leq y) \\ &= P\left(-\sqrt{y} \leq X \leq \sqrt{y}\right) \\ &= P(X \leq \sqrt{y}) - P(X < -\sqrt{y}) \\ &= P(X \leq \sqrt{y}) - P(X \leq -\sqrt{y}) \\ &= F_X\left(\sqrt{y}\right) - F_X\left(-\sqrt{y}\right) \end{aligned}$$

e, portanto, usando (3.1) e a regra da cadeia para cálculo de derivada de funções compostas, obtemos

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}\left[F_Y(y)\right] \\ &= \frac{d}{dy}\left[F_X\left(\sqrt{y}\right)\right] - \frac{d}{dy}\left[F_X\left(-\sqrt{y}\right)\right] \\ &= F_X'\left(\sqrt{y}\right)\frac{1}{2\sqrt{y}} - F_X'\left(-\sqrt{y}\right)\left(-\frac{1}{2\sqrt{y}}\right) \\ &= f_X\left(\sqrt{y}\right)\frac{1}{2\sqrt{y}} + f_X\left(-\sqrt{y}\right)\frac{1}{2\sqrt{y}} \end{aligned}$$

Como $0 \leq \sqrt{y} < 1$ e $-1 < -\sqrt{y} \leq 0$, resulta que $f_X\left(\sqrt{y}\right) = f_X\left(-\sqrt{y}\right) = \frac{1}{2}$. Logo

$$f_Y(y) = \begin{cases} \dfrac{1}{2\sqrt{y}} & \text{se } 0 \leq y < 1 \\ 0 & \text{caso contrário} \end{cases}$$

De modo análogo, para $0 \leq w < 1$

$$\begin{aligned} F_W(w) &= \Pr(W \leq w) = \Pr(|X| \leq w) = \Pr(-w \leq X \leq w) = \\ &= F_X(w) - F_X(-w) \end{aligned}$$

e, portanto

$$\begin{aligned} f_W(w) &= F_W'(w) = F_X'(w) - F_X'(-w)(-1) = \\ &= f_X(w) + f_X(-w) \end{aligned}$$

Como $0 \leq w < 1$ e $-1 < -w \leq 0$, resulta que $f_X(w) = f_X(-w) = \frac{1}{2}$. Logo

$$f_W(w) = \begin{cases} 1 & \text{se } 0 \leq y < 1 \\ 0 & \text{caso contrário} \end{cases}$$

que é a densidade uniforme padrão.

♦♦

3.2 Funções inversíveis

EXEMPLO 3.2 Transformação da uniforme

Sejam $X \sim Unif(0,1)$ e $Y = -\ln X$. Vamos calcular a função densidade de probabilidade de Y.

Solução

A função densidade de de X é

$$f_X(x) = \begin{cases} 1 & \text{se } 0 < x < 1 \\ 0 & \text{se } x \leq 0 \text{ ou } x \geq 1 \end{cases}$$

e sua função de distribuição é

$$F_X(x) = \begin{cases} 0 & \text{se } x < 0 \\ x & \text{se } 0 < x < 1 \\ 1 & \text{se } x \geq 1 \end{cases} \tag{3.2}$$

Vamos calcular a função de distribuição de $Y = g(X) = -\ln X$, observando algumas propriedades da função $y = g(x) = -\ln x$, cujo gráfico é dado na Figura 3.1. Tal função está definida para todo $x > 0$ e para esse domínio, sua imagem é o conjunto dos reais.

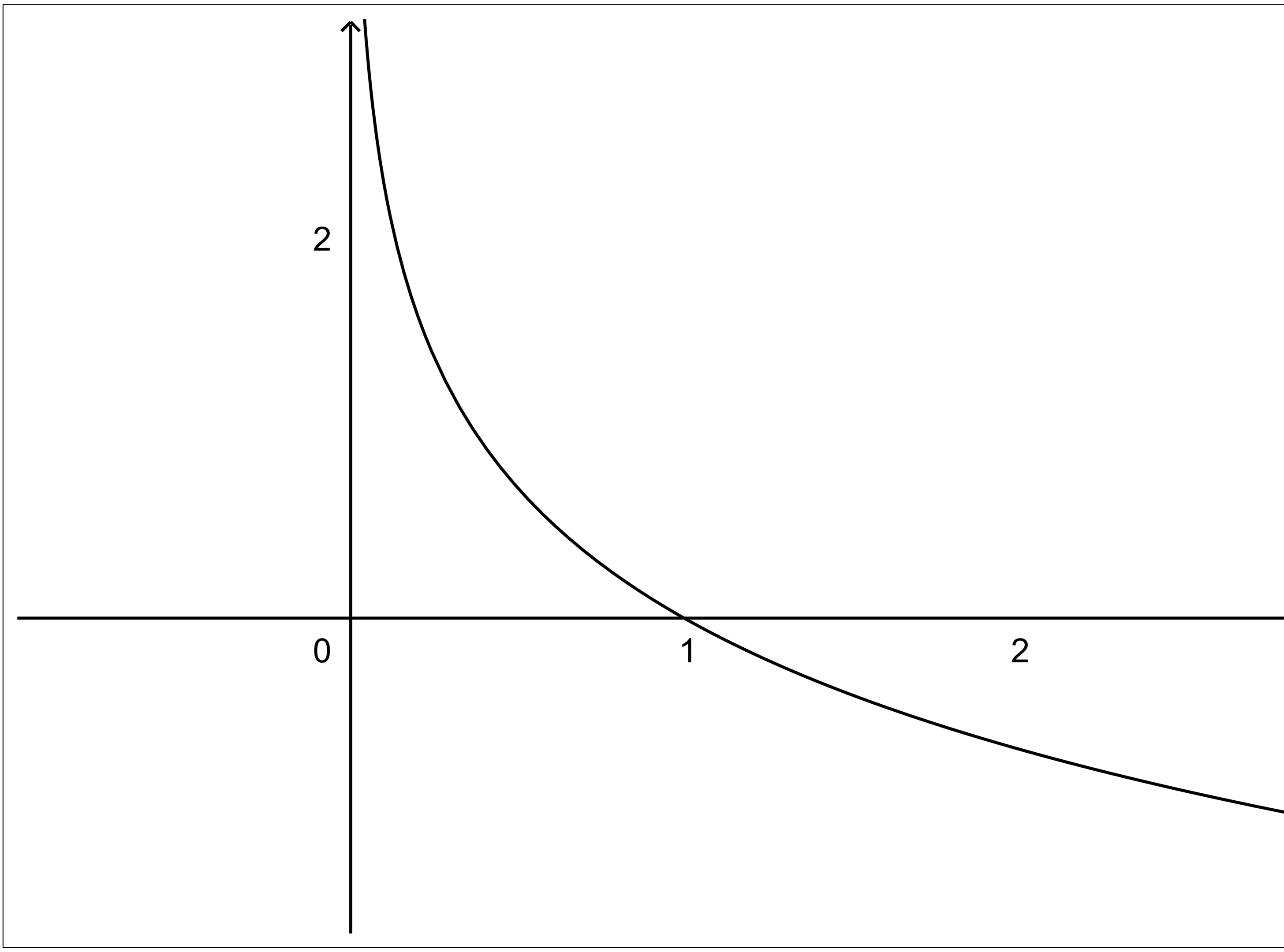


Figura 3.1 – Gráfico da função $y = g(x) = -\ln x$

Como esta é uma função estritamente decrescente, ela é inversível e sua inversa é calculada da seguinte maneira:

$$y = g(x) = -\ln x \Rightarrow \ln x = -y \Rightarrow g^{-1}(y) = x = e^{-y} \qquad y \in \mathbb{R} \tag{3.3}$$

Da definição de função estritamente crescente, resulta que (veja a Figura 3.2):

$$g(x) \leq y \Leftrightarrow x \geq g^{-1}(y) \tag{3.4}$$

Em termos das variáveis aleatórias X e Y, os valores possíveis de X estão no intervalo $(0, 1)$ (segmento em azul na Figura 3.2), e os valores possíveis de Y estão em $\mathbb{R}^+$. Veja a curva em vermelho na Figura 3.2.

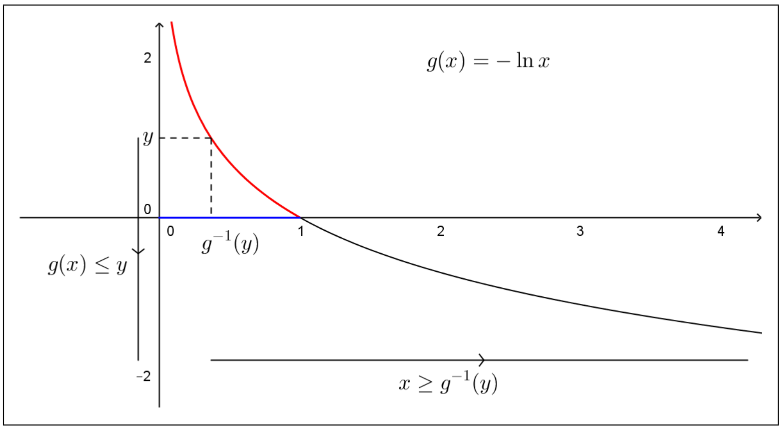


Figura 3.2 – Inversa de uma função decrescente

Usando (3.4), temos que, para $y > 0$

$$F_Y(y) = \mathrm{P}(Y \leq y) = \mathrm{P}[X \geq g^{-1}(y)] = 1 - \mathrm{P}[X < g^{-1}(y)] \Rightarrow$$

$$F_Y(y) \underbrace{=}_{X \text{ contínua}} 1 - F_X[g^{-1}(y)] \underbrace{=}_{(3.2)} 1 - g^{-1}(y) \underbrace{=}_{(3.3)} 1 - e^{-y} \tag{3.5}$$

Logo,

$$F_Y(y) = \begin{cases} 0 & y \leq 0 \\ 1 - e^{-y} & y > 0 \end{cases}$$

e

$$f_Y(y) = \frac{\mathrm{d}F_Y(y)}{\mathrm{d}y} = \begin{cases} 0 & y \leq 0 \\ e^{-y} & y > 0 \end{cases}$$

Note que essa é a densidade exponencial com parâmetro igual a 1.

EXEMPLO 3.3 Transformação linear

Seja X uma variável aleatória contínua com função densidade f_X e função de distribuição F_X. Se $Y = g(X) = aX + b$ em que $a \neq 0$ e b são constantes, determine a função densidade de Y.

Solução

Vamos calcular a função de distribuição de $Y = g(X) = aX + b$ observando propriedades da função $y = g(x) = ax + b$, que é uma função linear e, portanto, inversível, com inversa dada por

$$y = g(x) = ax + b \Leftrightarrow x = g^{-1}(y) = \frac{y - b}{a} \tag{3.6}$$

O sinal da inclinação depende do valor de a. Vamos analisar os dois casos, ilustrados nas Figuras 3.3 e 3.4, em que $a > 0$ e $a < 0$, respectivamente.

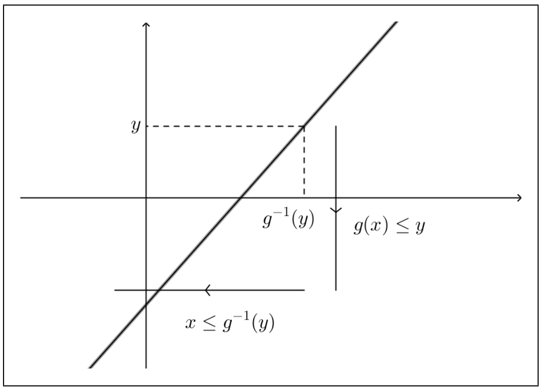


Figura 3.3 – $y = g(x) = ax + b, \;\; a > 0$

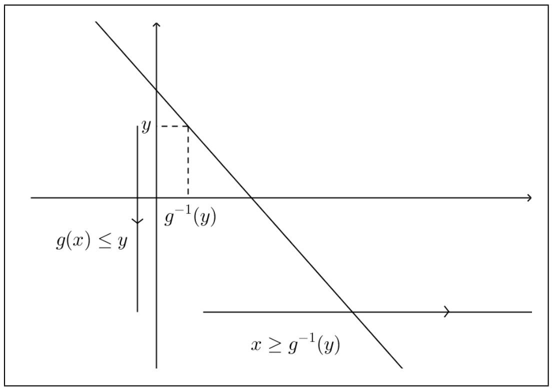


Figura 3.4 – $y = g(x) = ax + b, \;\; a < 0$

- $a > 0$ – função linear crescente

  Da Figura 3.3 podemos ver que

$$g(x) \leq y \Leftrightarrow x \leq g^{-1}(y)$$

Em termos das variáveis aleatórias X e Y, usando (3.6), temos que

$$F_Y(y) \;=\; \mathrm{P}(Y \leq y) = \mathrm{P}[g(X) \leq y] = \mathrm{P}[X \leq g^{-1}(y)] = \mathrm{P}\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right)$$

Logo,

$$f_Y(y) \;=\; F'_Y(y) = \frac{1}{a}F'_X\left(\frac{y-b}{a}\right) = \frac{1}{a}f_X\left(\frac{y-b}{a}\right) = \frac{1}{|a|}f_X\left(\frac{y-b}{a}\right) \tag{3.7}$$

- $a < 0$ – função linear decrescente

  Da Figura 3.4 podemos ver que

$$g(x) \leq y \Leftrightarrow x \geq g^{-1}(y)$$

Em termos das variáveis aleatórias X e Y, usando (3.6), temos que

$$\begin{aligned} F_Y(y) &= \mathrm{P}(Y \leq y) = \mathrm{P}[g(X) \leq y] = \mathrm{P}[X \geq g^{-1}(y)] = 1 - \mathrm{P}[X < g^{-1}(y)] \\ &= 1 - F_X\left[g^{-1}(y)\right] = 1 - F_X\left(\frac{y-b}{a}\right) \end{aligned}$$

Como $a < 0$, resulta que $-a > 0$, ou seja, $-a = |a|$. Logo,

$$f_Y(y) \;=\; F'_Y(y) = -\frac{1}{a}F'_X\left(\frac{y-b}{a}\right) = -\frac{1}{a}f_X\left(\frac{y-b}{a}\right) = \frac{1}{|a|}f_X\left(\frac{y-b}{a}\right) \tag{3.8}$$

Os dois resultados (3.7) e (3.8) constituem a prova do

TEOREMA 3.1

Seja X uma variável aleatória contínua com função densidade $f_X(x)$. Se $Y = aX + b$, $a \neq 0$ e b são constantes, então

$$f_Y(y) = \frac{1}{|a|}f_X\left(\frac{y-b}{a}\right) \tag{3.9}$$

▼

EXEMPLO 3.4 Mais sobre transformação linear

Se a função densidade da variável aleatória X é dada por

$$f(x) = \begin{cases} 3x^2 & -1 \leq x \leq 0 \\ 0 & x < -1 \text{ ou } x > 0 \end{cases}$$

calcule a função densidade de $Y = 2X - \frac{3}{5}$, bem como sua esperança e sua variância.

Solução

Temos aqui uma transformação linear crescente com $a = 2$ e $b = -0,6$. Como $-1 \leq x \leq 0$, resulta que $-2,6 \leq y \leq -0,6$. Logo,

$$f_Y(y) = f_X\left(\frac{y+0,6}{2}\right) \times \frac{1}{2} \qquad \text{se } -2,6 \leq y \leq -0,6$$

ou seja

$$f_Y(y) = 3\left(\frac{y+0,6}{2}\right)^2 \times \frac{1}{2} = \frac{3}{8}(y+0,6)^2 \qquad \text{se } -2,6 \leq y \leq -0,6$$

Pelas propriedades da esperança e da variância, se $Y = 2X - \frac{3}{5}$ então

$$\begin{aligned} \mathrm{E}(Y) &= 2\,\mathrm{E}(X) - \frac{3}{5} \\ \mathrm{Var}(Y) &= 4\,\mathrm{Var}(X) \end{aligned}$$

$$\mathrm{E}(X) = \int_{-1}^{0} x3x^2 dx = \left(3\frac{x^4}{4}\right)\Big|_{-1}^{0} = -\frac{3}{4} \Longrightarrow E(Y) = -\frac{6}{4} - \frac{3}{5} = \frac{-30-12}{20} = -2,1$$

$$\begin{aligned} \mathrm{E}(X^2) &= \int_{-1}^{0} x^2 3x^2 dx = \left(3\frac{x^5}{5}\right)\Big|_{-1}^{0} = \frac{3}{5} \Longrightarrow Var(X) = \frac{3}{5} - \left(-\frac{3}{4}\right)^2 = \frac{48-45}{80} = \frac{3}{80} \\ &\Longrightarrow \quad Var(Y) = 4 \times \frac{3}{80} = \frac{3}{20} = 0,15 \end{aligned}$$

Podemos, também, calcular $\mathrm{E}(Y)$ e $\mathrm{E}^2(Y)$ a partir da função densidade de Y:

$$\begin{aligned} \mathrm{E}(Y) &= \int_{-2,6}^{-0,6} y\frac{3}{8}(y+0,6)^2\,\mathrm{d}y = \int_{-2,6}^{-0,6} \frac{3}{8}(y^3 + 1,2y^2 + 0,36y)\mathrm{d}y = \frac{3}{8}\left[\frac{y^4}{4} + 1,2\frac{y^3}{3} + 0,36\frac{y^2}{2}\right]_{-2,6}^{-0,6} \\ &= \frac{3}{8}\left[\frac{(-0,6)^4}{4} + 0,4(-0,6)^3 + 0,18(-0,6)^2\right] - \frac{3}{8}\left[\frac{(-2,6)^4}{4} + 0,4(-2,6)^3 + 0,18(-2,6)^2\right] \\ &= \frac{3}{8}[(0,0324 - 0,0864 + 0,0648) - (11,4244 - 7,0304 + 1,2168)] = \frac{3}{8}(0,0108 - 5,6108) = -2,1 \end{aligned}$$

$$\begin{aligned} \mathrm{E}(Y^2) &= \int_{-2,6}^{-0,6} y^2\frac{3}{8}(y+0,6)^2\,\mathrm{d}y = \int_{-2,6}^{-0,6} \frac{3}{8}(y^4 + 1,2y^3 + 0,36y^2)\mathrm{d}y = \frac{3}{8}\left[\frac{y^5}{5} + 1,2\frac{y^4}{4} + 0,36\frac{y^3}{3}\right]_{-2,6}^{-0,6} \\ &= \frac{3}{8}\left[\frac{(-0,6)^5}{5} + 0,3(-0,6)^4 + 0,12(-0,6)^3\right] - \frac{3}{8}\left[\frac{(-2,6)^5}{5} + 0,3(-2,6)^4 + 0,12(-2,6)^3\right] \\ &= \frac{3}{8}[-0,002592 - (-12,162592)] = 4,56 \end{aligned}$$

Logo, $\text{Var}(Y) = 4,56 - 2,1^2 = 0,15$.

Capítulo 4

Sobre a Distribuição Normal

Neste capítulo estudaremos uma das principais distribuiç˜es contínuas.

4.1 Alguns resultados de Cálculo

Com o uso de coordenadas polares, pode-se mostrar que

$$\int_0^\infty \exp\left(-\frac{t^2}{2}\right) dt = \sqrt{\frac{\pi}{2}} \tag{4.1}$$

Como o integrando é uma função par, temos também que

$$\int_{-\infty}^\infty \exp\left(-\frac{t^2}{2}\right) dt = 2 \times \int_0^\infty \exp\left(-\frac{t^2}{2}\right) dt = 2 \times \sqrt{\frac{\pi}{2}} = \sqrt{2\pi}$$

ou ainda

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \exp\left(-\frac{t^2}{2}\right) dt = 1 \tag{4.2}$$

EXEMPLO 4.1 Cálculo de de $\Gamma\left(\frac{1}{2}\right)$

Da definição da função gama, temos que

$$\Gamma(1/2) = \int_0^\infty e^{-x} x^{1/2-1} dx = \int_0^\infty \frac{e^{-x}}{\sqrt{x}} dx$$

Considere a seguinte transformação de variável:

$$x = \frac{t^2}{2}$$

Então,

$$\begin{array}{rcl} dx & = & tdt \\ x & = & 0 \Rightarrow t = 0 \\ x & \to & \infty \Rightarrow t \to \infty \end{array}$$

Logo,

$$\Gamma(1/2) = \int_0^\infty \left(\frac{e^{-t^2/2}}{\sqrt{\frac{t^2}{2}}} \right) t dt = \sqrt{2} \int_0^\infty e^{-t^2/2} dt = \sqrt{2} \times \sqrt{\frac{\pi}{2}}$$

ou seja:

$$\Gamma(1/2) = \sqrt{\pi} \tag{4.3}$$

4.2 Densidade normal padrão

4.2.1 Definição

Analisando a equação (4.2), vemos que a função $\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$ satisfaz as condições para ser uma função densidade.

DEFINIÇÃO Densidade normal padrão

A variável aleatória Z tem distribuição normal padrão se sua função densidade é dada por

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \tag{4.4}$$

para todo $z \in \mathbb{R}$. Vamos denotar por $N(0;1)$ a densidade normal padrão e, se uma variável aleatória Z é distribuída segundo uma normal padrão, representaremos esse fato como $Z \sim N(0;1)$.

4.2.2 Esperança e variância

Considere a expressão da densidade normal padrão dada em (4.4). Podemos ver que o expoente de x é 2, o que significa que $\varphi(z) = \varphi(-z)$, ou seja, $\varphi(z)$ é simétrica em torno de 0. Resulta, então, que

$$Z \sim N(0;1) \Longrightarrow \mathrm{E}(Z) = 0$$

Como $\mathrm{E}(Z) = 0$ então

$$\mathrm{Var}(Z) = \mathrm{E}(Z^2) = \int_{-\infty}^{+\infty} x^2 \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 \exp\left(-\frac{x^2}{2}\right) dx$$

Note que o integrando é uma função simétrica em torno de 0 (veja a Figura 4.1). Resulta, então, que (note os limites de integração)

$$\mathrm{Var}(Z) = \mathrm{E}(Z^2) = 2 \times \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} x^2 \exp\left(-\frac{x^2}{2}\right) dx$$

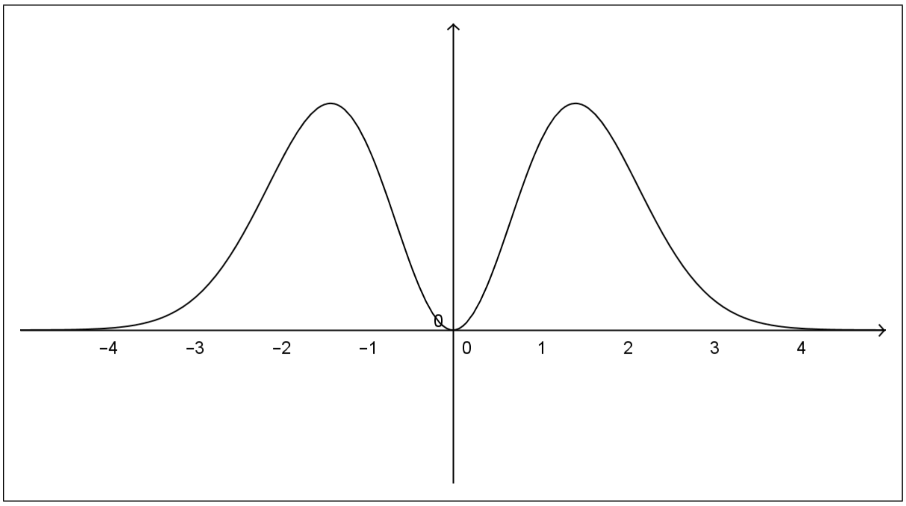


Figura 4.1 – Gráfico de $f(x) = x^2 \exp\left(-\frac{x^2}{2}\right)$

Esta integral é calculada usando-se o método de integração por partes. Fazendo:

- $x \exp\left(-\frac{x^2}{2}\right) dx = dv \Rightarrow v = -\exp\left(-\frac{x^2}{2}\right)$
- $x = u \Rightarrow dx = du$

obtemos

$$-x \exp\left(-\frac{x^2}{2}\right)\Bigg|_0^\infty = \int_0^\infty \left[-\exp\left(-\frac{x^2}{2}\right)\right] dx + \int_0^\infty x^2 \exp\left(-\frac{x^2}{2}\right) dx \tag{4.5}$$

Pelos resultados (2.5) e (4.1) resulta

$$0 = -\sqrt{\frac{\pi}{2}} + \int_0^\infty x^2 \exp\left(-\frac{x^2}{2}\right) dx \Longrightarrow \int_0^\infty x^2 \exp\left(-\frac{x^2}{2}\right) dx = \sqrt{\frac{\pi}{2}}$$

Logo,

$$\text{Var}(Z) = \frac{2}{\sqrt{2\pi}} \times \sqrt{\frac{\pi}{2}} \Rightarrow \text{Var}(Z) = 1 \tag{4.6}$$

Resumindo:

$$Z \sim N(0;1) \quad \Rightarrow \quad \begin{cases} \text{E}(Z) = 0 \\ \text{Var}(Z) = 1 \end{cases} \tag{4.7}$$

4.2.3 Características da curva normal padrão

1. Simétrica em torno de 0; note que $\varphi(-x) = \varphi(x)$.
2. Assíntotas: $\lim\limits_{x \to -\infty} \varphi(x) = \lim\limits_{x \to \infty} \varphi(x) = 0$; esse resultado segue diretamente do fato de que $\lim_{x\to\infty} e^{-x} = 0$
3. Ponto de máximo

Para calcular a primeira e segunda derivadas de $\varphi(x)$, devemos lembrar que $(e^x)' = e^x$ e, pela regra da cadeia, $(e^{g(x)})' = e^{g(x)}g'(x)$. Aplicando esses resultados à densidade normal padrão, obtemos que:

$$\varphi'(x) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{x^2}{2}\right]\left[-\frac{1}{2}2x\right] = -\varphi(x)x \tag{4.8}$$

Derivando novamente, obtemos:

$$\begin{aligned} \varphi''(x) &= -\varphi'(x)x - \varphi(x) = -[-\varphi(x)x]x - \varphi(x) \\ &= \varphi(x)x^2 - \varphi(x) = \varphi(x)(x^2 - 1) \end{aligned} \tag{4.9}$$

Analisando a equação (4.8) e lembrando que $\varphi(x) > 0$, pode-se ver que:

$$\varphi'(x) = 0 \Leftrightarrow x = 0$$

e assim, $x = 0$ é um ponto crítico. Como $\varphi'(x) > 0$ para $x < 0$ e $\varphi'(x) < 0$ para $x > 0$, então φ é crescente à esquerda de 0 e decrescente à direita de 0. Segue, então, que $x = 0$ é um ponto de máximo e nesse ponto

$$\varphi(0) = \frac{1}{\sqrt{2\pi}} \tag{4.10}$$

4. Pontos de inflexão

Analisando a segunda derivada dada por (4.9), tem-se que:

$$\varphi''(x) = 0 \Leftrightarrow x^2 - 1 = 0 \Leftrightarrow x = \pm 1 \tag{4.11}$$

Além disso,

$$\varphi''(x) > 0 \Leftrightarrow x^2 - 1 > 0 \Leftrightarrow x^2 > 1 \Leftrightarrow x > 1 \quad \text{ou} \qquad x < -1$$

e

$$\varphi''(x) < 0 \Leftrightarrow x^2 - 1 < 0 \Leftrightarrow x^2 < 1 \Leftrightarrow -1 < x < 1$$

Logo, $\varphi(x)$ é côncava para cima se $x > 1$ ou $x < -1$ e é côncava para baixo quando $-1 < x < +1$.

Na Figura 4.2 temos o gráfico da densidade normal padrão; aí as linhas pontilhadas indicam a ocorrência dos pontos de inflexão

4.2.4 Função de distribuição acumulada

A função de distribuição acumulada de qualquer variável aleatória X é definida por $F_X(X) = \Pr(X \leq x)$. No caso da densidade normal padrão, essa função é dada pela integral

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) dt \tag{4.12}$$

para a qual não existe uma antiderivada em forma de função elementar. Assim, a função de distribuição acumulada da normal padrão é calculada por integração numérica. Todos os pacotes estatísticos possuem rotinas especiais para esse cálculo. No EXCEL, a função DIST.NORMP calcula $\mathrm{P}(Z \leq x)$ para qualquer x, onde $Z \sim N(0;1)$.

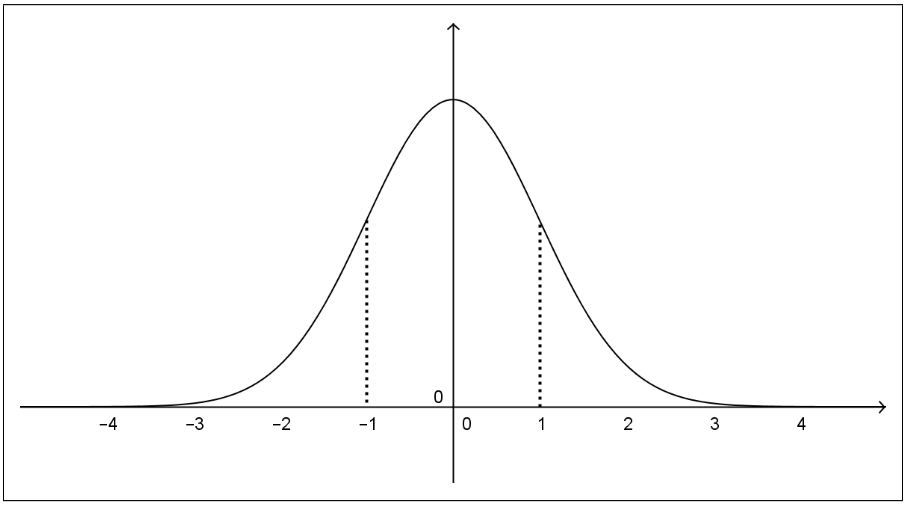


Figura 4.2 – Densidade normal padrão

4.2.5 A Tabela da Normal Padrão

Vimos anteriormente que o cálculo de probabilidades associadas a variáveis aleatórias contínuas envolve cálculo de áreas sob a curva densidade (mais precisamente, cálculo de integral da fdp). Isso, obviamente, continua valendo para a densidade normal. A diferença está no fato de que o cálculo de áreas sob a curva normal envolve métodos numéricos mais complexos e, para facilitar esses cálculos, podemos usar uma tabela em que alguns valores já se encontram calculados.

Este curso terá como base a Tabela 1 apresentada no Apêndice ??, embora muitos livros utilizem a tabela da distribuição acumulada dada na Tabela 2 do mesmo apêndice, que discutiremos no final desta seção.

A Tabela 1 será usada para calcular probabilidades associadas a uma variável aleatória normal padrão Z. Assim, com essa tabela, poderemos calcular probabilidades do tipo $\text{P}(Z > 1), \text{P}(Z \leq 3), \text{P}(-1 \leq Z \leq 2)$ etc.

Vamos analisar cuidadosamente esta tabela. A partir do cabeçalho e do gráfico na tabela, podemos ver que as entradas no corpo da tabela fornecem probabilidades do tipo $\text{P}(0 \leq Z \leq z)$. Com relação à abcissa z, seus valores são apresentados na tabela ao longo da coluna lateral à esquerda em conjunto com a linha superior, ambas sombreadas de cinza. Na coluna à esquerda, temos a casa inteira e a primeira casa decimal; na linha superior, temos a segunda casa decimal. Por exemplo, ao longo da primeira linha da tabela, temos probabilidades associadas às abcissas 0,00; 0,01; 0,02, ..., 0,09; na segunda linha da tabela, temos probabilidades associadas às abcissas 0,10; 0,11; 0,12; ..., 0,19; na última linha da tabela, temos probabilidades associadas às abcissas 4,00; 4,01; 4,02; ...; 4,09.

A entrada 0,00000 no canto superior esquerdo da tabela corresponde à seguinte probabilidade: $\text{P}(0 \leq Z \leq 0,00)$, ou seja, $\text{P}(Z = 0)$ e, como visto, essa probabilidade é nula, uma vez que, para qualquer variável aleatória contínua X, $\text{P}(X = x_0) = 0$. A segunda entrada na primeira linha, 0,00399, corresponde a $\text{P}(0 \leq Z \leq 0,01)$, que é a área sob a curva densidade normal padronizada compreendida entre os valores 0 e 0,01 (veja o gráfico na tabela).

Note que esta tabela apresenta probabilidades correspondentes a abcissas positivas, ou seja, esta tabela trata de área sob a curva no lado positivo do eixo. Para calcular áreas no lado negativo, teremos de usar o fato de a curva da densidade normal ser simétrica. Sempre faça um esboço da curva densidade, sombreando a área correspondente à probabilidade desejada; isso lhe ajudará no cálculo da probabildiade. Vamos terminar esta seção apresentando vários exemplos de cálculos de probabilidades de uma v.a. Z com distribuição normal padrão, ou seja, no que segue, $Z \sim N(0;1)$. Os exemplos apresentados cobrem todas as situações possíveis. Assim, é importante que você entenda bem a situação ilustrada por cada um dos exemplos, para poder aplicar o método de solução adequano aos novos exercícios.

Para simplificar a solução dos exercícios, vamos adotar a seguinte notação.

! Entradas da Tabela 1 do Apêndice ??

Vamos representar por $tab(z)$ as entradas da Tabela 1 do Apêndice ??, ou seja,

$$tab(z) = \mathrm{P}(0 \leq Z \leq z)$$

EXEMPLO 4.2 $\mathrm{P}(0 \leq Z \leq 1,22)$

Veja as Figuras 4.3 e 4.4. Essa probabilidade é dada diretamente na Tabela 1, utilizando a entrada correspondente à linha 1,2 e à coluna com o valor 2. O resultado é

$$\mathrm{P}(0 \leq Z \leq 1,22) = tab(1,22) = 0,3888$$

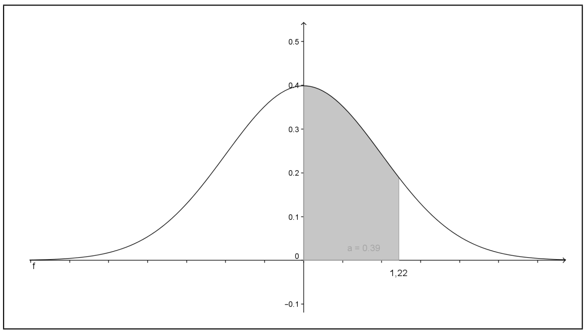


Figura 4.3 – $\mathrm{P}(0 \leq Z \leq 1,22)$ como área

e 1ª. Decimal	0	1	2	3
0,0	0,0000	0,0040	0,0080	0,0120
0,1	0,0398	0,0438	0,0478	0,0517
0,9	0,3159	0,3186	0,3212	0,3238
1,0	0,3413	0,3438	0,3461	0,3485
1,1	0,3643	0,3665	0,3686	0,3708
1,2	0,3849	0,3869	**0,3888**	0,3907
1,3	0,4032	0,4049	0,4066	0,4082

Figura 4.4 – $\mathrm{P}(0 \leq Z \leq 1,22)$ - Uso da tabela

♦♦

EXEMPLO 4.3 $\mathrm{P}(1 \leq Z \leq 2)$

Note que este exemplo trata da área (probabilidade) entre duas abcissas positivas. Na Figura 4.5 ilustra-se a área (probabilidade) desejada. Note que esta área pode ser obtida pela diferença entre as áreas das Figuras 4.6 e 4.8, cujos valores são encontrados na Tabela 1 conforme ilustram as Figuras 4.7 e 4.9.

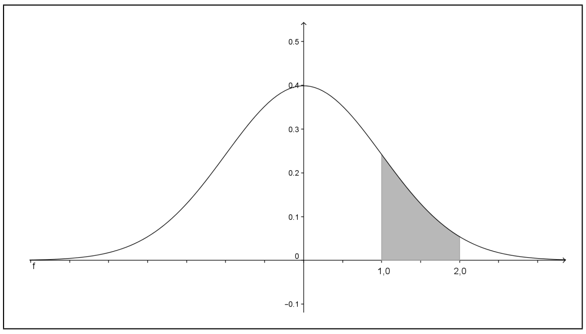


Figura 4.5 – $P(1 \leq Z \leq 2)$ como área

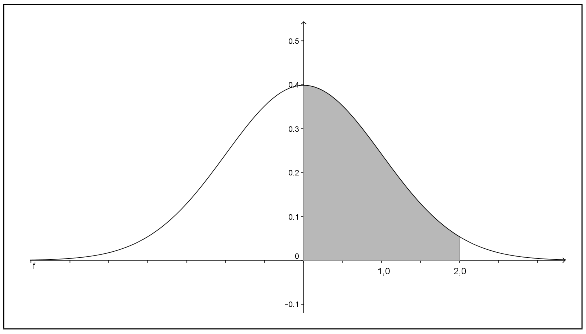


Figura 4.6 – $P(0 \leq Z \leq 2)$ como área

e 1ª. Decimal	0	1	2	3	4
0,0	0,0000	0,0040	0,0080	0,0120	0,0160
0,1	0,0398	0,0438	0,0478	0,0517	0,0557
0,2	0,0793	0,0832	0,0871	0,0910	0,0948
1,8	0,4641	0,4649	0,4656	0,4664	0,4671
1,9	0,4713	0,4719	0,4726	0,4732	0,4738
2,0	**0,4772**	0,4778	0,4783	0,4788	0,4793
2,1	0,4821	0,4826	0,4830	0,4834	0,4838

Figura 4.7 – $P(0 \leq Z \leq 2)$ - Uso da tabela

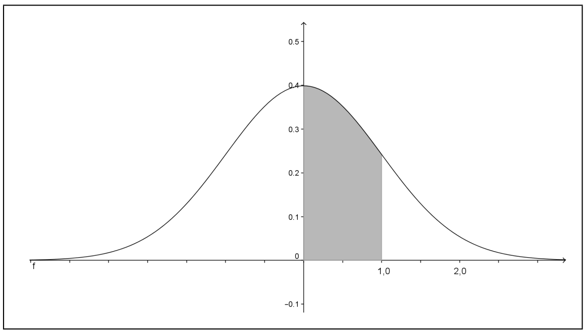


Figura 4.8 – $P(0 \leq Z \leq 1)$ como área

e 1ª. Decimal	0	1	2	3	4
0,0	0,0000	0,0040	0,0080	0,0120	0,0160
0,1	0,0398	0,0438	0,0478	0,0517	0,0557
0,2	0,0793	0,0832	0,0871	0,0910	0,0948
0,8	0,2881	0,2910	0,2939	0,2967	0,2995
0,9	0,3159	0,3186	0,3212	0,3238	0,3264
1,0	**0,3413**	0,3438	0,3461	0,3485	0,3508
1,1	0,3643	0,3665	0,3686	0,3708	0,3729

Figura 4.9 – $P(0 \leq Z \leq 1)$ - Uso da tabela

Concluimos, então, que

$$P(1 \leq Z \leq 2) = P(0 \leq Z \leq 2) - P(0 \leq Z \leq 1) = tab(2,0) - tab(1,0) = 0,4772 - 0,3413 = 0,1359$$

♦♦

EXEMPLO 4.4 $P(Z \geq 1)$

Note que este exemplo trata da área (probabilidade) à direita (≥) de uma abcissa positiva. Na Figura 4.10 ilustra-se a área (probabilidade) desejada. Note que esta área pode ser obtida pela diferença entre as áreas das Figuras 4.11 e 4.12. A primeira área corresponde à probabilidade $P(Z \geq 0)$ e é igual a 0,5, pois a média $\mu = 0$ é o eixo de simetria e a área total é 1. Logo, $P(Z \geq 0) = P(Z \leq 0) = 0,5$. A segunda área vem direto da Tabela 1.

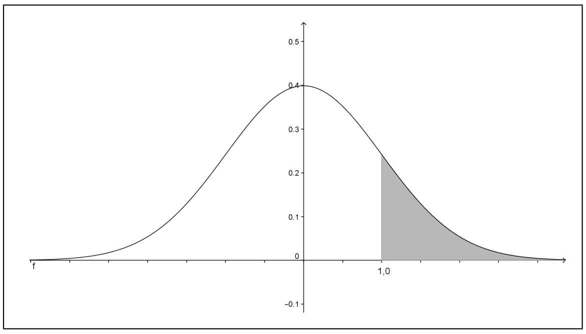


Figura 4.10 – $\text{P}(Z \geq 1)$

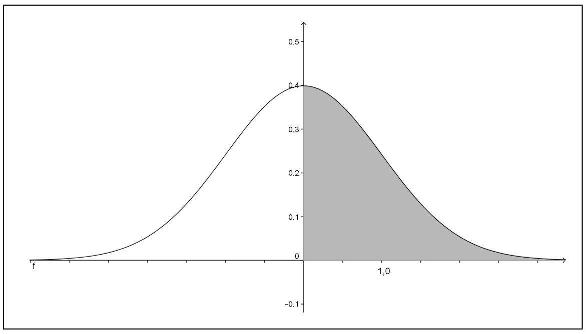


Figura 4.11 – $\text{P}(Z \geq 0)$

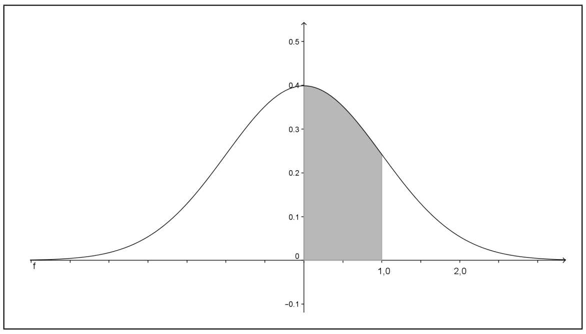


Figura 4.12 – $\text{P}(0 \leq Z \leq 1)$

Concluimos, então, que

$$\text{P}(Z \geq 1) = \text{P}(Z \geq 0) - \text{P}(0 \leq Z \leq 1) = 0,5 - tab(1,0) = 0,5 - 0,3413 = 0,1587$$

♦♦

EXEMPLO 4.5 $\text{P}(Z \leq 1)$

Note que este exemplo trata da área (probabilidade) à esquerda (≤) de uma abcissa positiva. Na Figura 4.13 ilustra-se a área (probabilidade) desejada. Note que esta área pode ser obtida pela soma das áreas das Figuras 4.14 e 4.15. A primeira área corresponde à probabilidade $\text{P}(Z \leq 0)$ e é igual a 0,5, conforme visto no exemplo anterior.

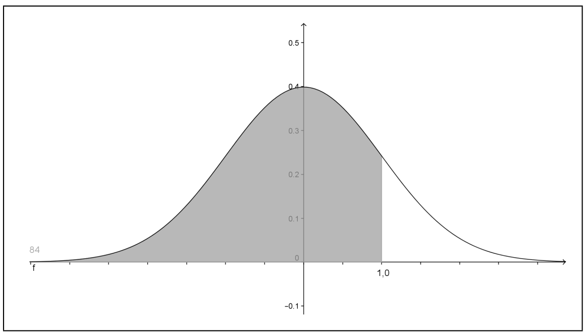


Figura 4.13 – $\text{P}(Z \leq 1)$

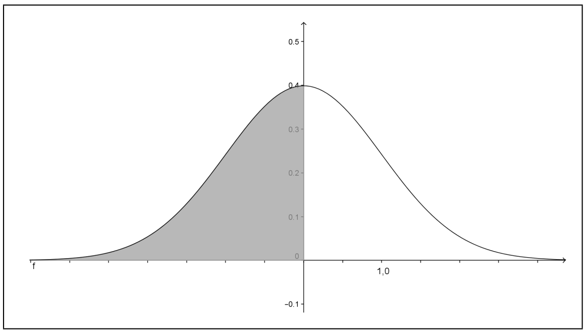


Figura 4.14 – $\mathrm{P}(Z \leq 0)$

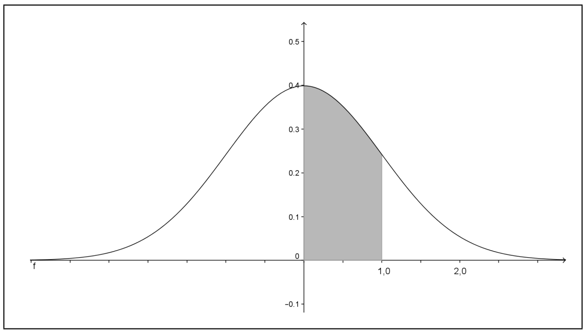


Figura 4.15 – $\mathrm{P}(0 \leq Z \leq 1)$

Concluimos, então, que

$$\mathrm{P}(Z \geq 1) = \mathrm{P}(Z \geq 0) + \mathrm{P}(0 \leq Z \leq 1) = 0,5 + tab(1,0) = 0,5 + 0,3413 = 0,8413$$

♦♦

EXEMPLO 4.6 $\mathrm{P}(Z \leq -0,5)$

Note que este exemplo trata da área (probabilidade) à esquerda (≤) de uma abcissa negativa e, agora, começamos a trabalhar com abcissas negativas. Na Figura 4.16 ilustra-se a área (probabilidade) desejada. Pela simetria da curva densidade normal, resulta que essa área é igual à área ilustrada na Figura 4.17.

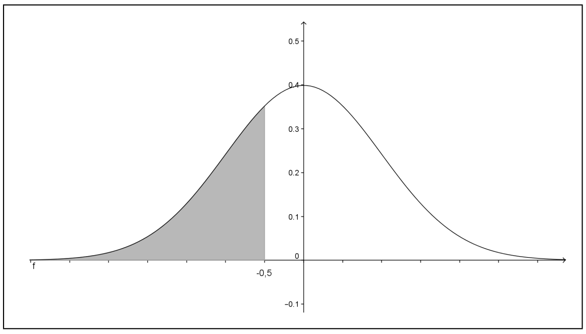


Figura 4.16 – $\mathrm{P}(Z \leq -0,5)$

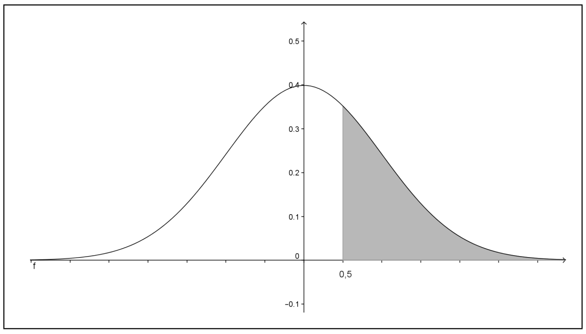


Figura 4.17 – $\mathrm{P}(Z \geq 0,5)$

Concluimos, então, que

$$\mathrm{P}(Z \leq -0,5) = \mathrm{P}(Z \geq 0,5) = 0,5 - \mathrm{P}(0 \leq Z < 0,5) = 0,5 - tab(0,5) = 0,5 - 0,1915 = 0,3085$$

♦♦

EXEMPLO 4.7 $\mathrm{P}(Z \geq -0,5)$

Note que este exemplo trata da área (probabilidade) à direira (≥) de uma abcissa negativa. Na Figura 4.18 ilustra-se a área (probabilidade) desejada. Essa área é a soma das áreas representadas nas Figuras 4.19 e 4.20. Essa última área, por sua vez, é igual à area representada na Figura 4.21, pela simetria da curva densidade.

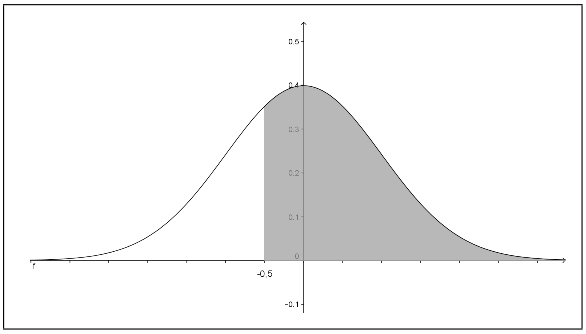


Figura 4.18 – $P(Z \geq -0,5)$

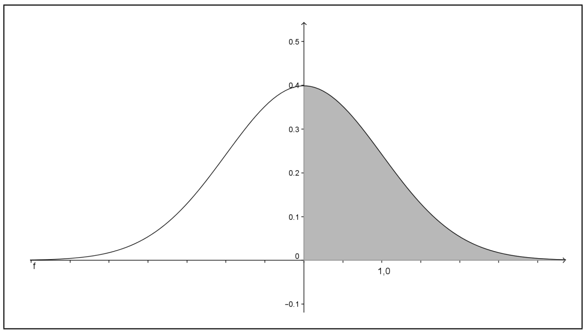


Figura 4.19 – $P(Z \geq 0)$

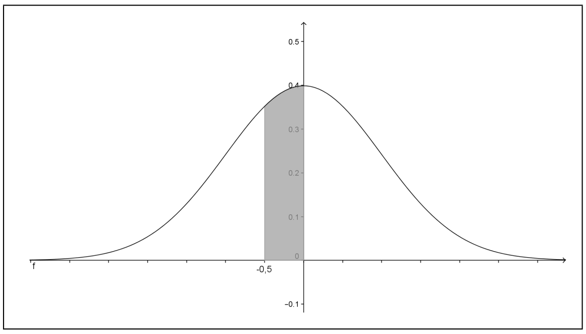


Figura 4.20 – $P(-0,5 \leq Z \leq 0)$

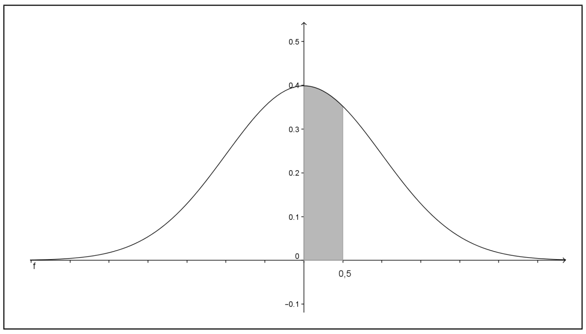


Figura 4.21 – $P(Z \geq 0,5)$

Concluimos, então, que

$$\begin{aligned} P(Z \geq -0,5) &= P(-0,5 \leq Z \leq 0) + P(Z \geq 0) \\ &= P(0 \leq Z < 0,5) + 0,5 = tab(0,5) + 0,5 = 0,1915 + 0,5 = 0,6915 \end{aligned}$$

♦♦

EXEMPLO 4.8 $P(-2,1 \leq Z \leq -1,4)$

Note que este exemplo trata da área (probabilidade) entre duas abcissas negativas. Na Figura 4.22 ilustra-se a área (probabilidade) desejada. Por simetria, essa área é igual à área ilustrada na Figura 4.23, já analisada no Exemplo 4.3.

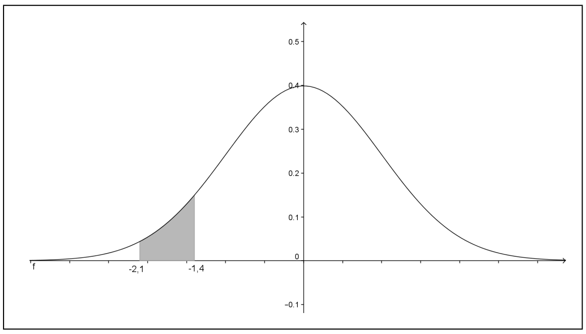


Figura 4.22 – $P(-2,1 \leq Z \leq -1,4)$

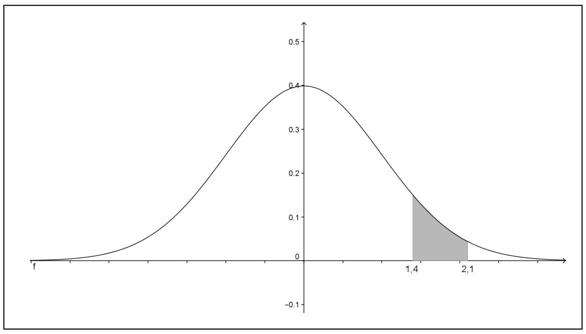


Figura 4.23 – $P(1,4 \leq Z \leq -2,1)$

Concluimos, então, que

$$\begin{aligned} P(-2,1 \leq Z \leq -1,4) &= P(1,4 \leq Z \leq 2,1) = P(0 \leq Z \leq 2,1) - P(0 \leq Z \leq 1,4) \\ &= tab(2,1) - tab(1,4) = 0,4821 - 0,4192 = 0,0629 \end{aligned}$$

♦♦

EXEMPLO 4.9 $P(-2,1 \leq Z \leq 1,4)$

Note que este exemplo trata da área (probabilidade) entre duas abcissas, uma negativa e outra positiva. Na Figura 4.24 ilustra-se a área (probabilidade) desejada. Essa área é a soma das áreas representadas nas Figuras 4.25 e 4.26. Por simetria, essa última área é igual à área sombreada na Figura 4.27, o que nos leva à cocnlusão de que

$$\begin{aligned} P(-2,1 \leq Z \leq 1,4) &= P(0 \leq Z \leq 1,4) + P(-2,1 \leq Z \leq 0) \\ &= P(0 \leq Z \leq 1,4) + P(0 \leq Z \leq 2,1) = tab(1,4) + tab(2,1) \\ &= 0,4821 + 0,4192 = 0,9013 \end{aligned}$$

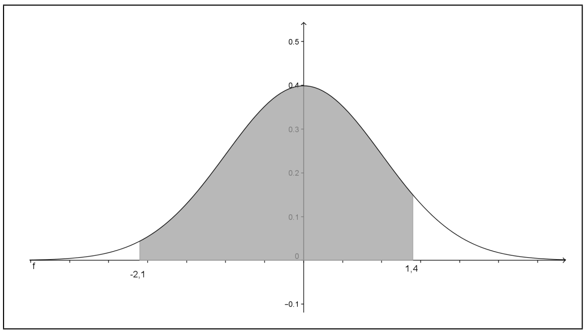


Figura 4.24 – $P(-2,1 \leq Z \leq 1,4)$

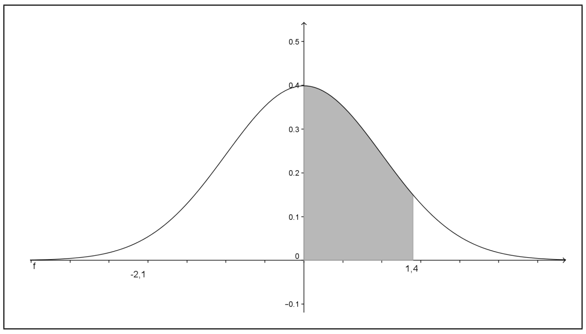


Figura 4.25 – $P(0 \leq Z \leq 1,4)$

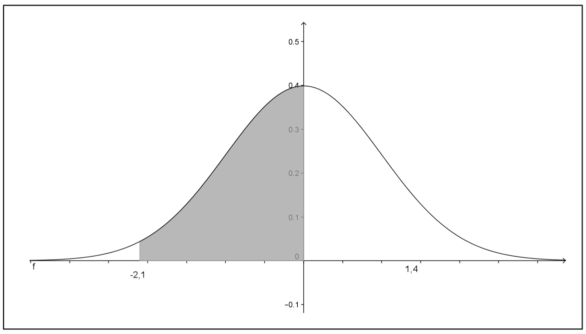


Figura 4.26 – $P(-2,1 \leq Z \leq 0)$

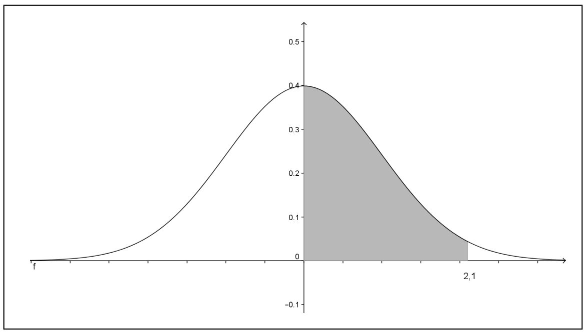


Figura 4.27 – $P(0 \leq Z \leq 2,1)$

4.2.6 A Tabela da Distribuição Acumulada da Normal Padrão

Muitos livros trabalham com a tabela da função de distribuição acumulada da normal padrão, que representaremos pela letra grega fi maiúscula, Φ:

$$\Phi(z) = P(Z \leq z).$$

A Tabela 2 do Apêndice **??** apresenta os valores de $\Phi(z)$ para $z \geq 0$. Vamos usar essa tabela para refazer os exemplos vistos anteriormente, que serão apresentados em uma ordem diferente, mais idaticamente apropriada para esse contexto.

EXEMPLO 4.10 $P(Z \leq 1)$

Essa probabilidade resulta diretamente da definição de distribuição acumulada:

$$P(Z \leq 1) = \Phi(1,0) = 0,8413$$

♦♦

EXEMPLO 4.11 $P(Z \geq 1)$

Pela lei do complementar, temos que

$$P(Z \geq 1) = 1 - P(Z < 1)$$

Mas, como Z é uma variável aleatória contínua, sabemos que $P(Z = z) = 0$. Logo

$$P(Z < z) = P(Z \leq z)$$

Logo,

$$P(Z \geq 1) = 1 - P(Z < 1) = 1 - P(Z \leq 1) = 1 - \Phi(1,0) = 1 - 0,8413 = 0,1587$$

♦♦

EXEMPLO 4.12 $P(Z \leq -0,5)$

Vimos, no Exemplo 4.6, que

$$P(Z \leq -0,5) = P(Z \geq 0,5)$$

Logo,

$$P(Z \leq -0,5) = P(Z \geq 0,5) = 1-P(Z < 0,5) = 1-P(Z \leq 0,5) = 1-\Phi(0,5) = 1-0,6915 = 0,3085$$

♦♦

EXEMPLO 4.13 $P(Z \geq -0,5)$

Veja as Figuras 4.28 e 4.29.

$$\begin{aligned} P(Z \geq -0,5) &= 1 - P(Z < -0,5) = 1 - P(Z > 0,5) = 1 - [1 - P(Z \leq 0,5)] \\ &= P(Z \leq 0,5) = \Phi(0,5) = 0,6915 \end{aligned}$$

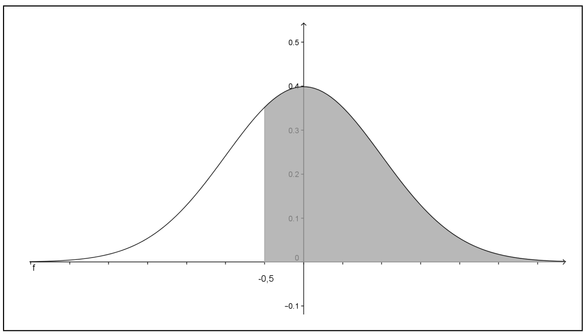


Figura 4.28 – $P(Z \geq -0,5)$

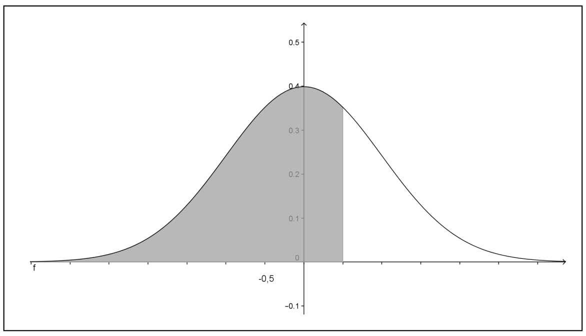


Figura 4.29 – $P(Z \leq 0,5)$

♦♦

EXEMPLO 4.14 $P(0 \leq Z \leq 1,22)$

Veja as Figuras 4.3, 4.30 e 4.31.

$$P(0 \leq Z \leq 1,22) = P(Z \leq 1,22) - P(Z \leq 0) = \Phi(1,22) - 0,5 = 0,8888 - 0,5 = 0,3888$$

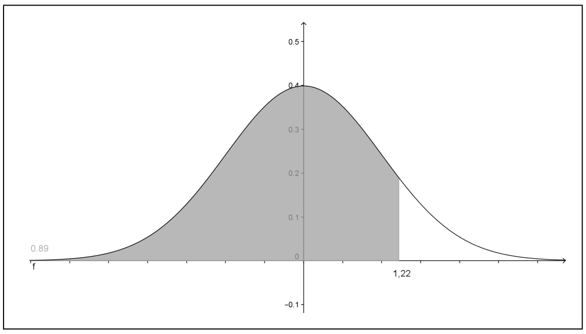


Figura 4.30 – $P(Z \leq 1,22)$

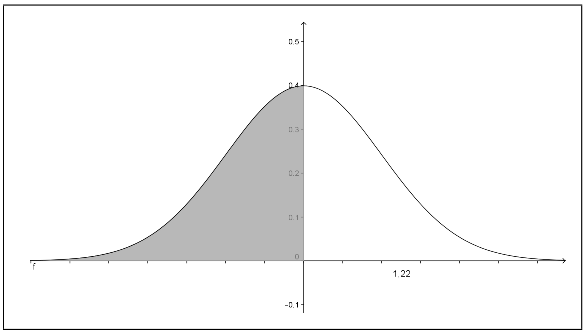


Figura 4.31 – $P(Z \leq 0)$

♦♦

EXEMPLO 4.15 $P(1 \leq Z \leq 2)$

Veja as Figuras 4.5, 4.32 e 4.33.

$$\begin{aligned} P(1 \leq Z \leq 2) &= P(Z \leq 2) - P(Z < 1) = P(Z \leq 2) - P(Z \leq 1) = \Phi(2,0) - \Phi(1,0) \\ &= 0,9772 - 0,8413 = 0,1359 \end{aligned}$$

♦♦

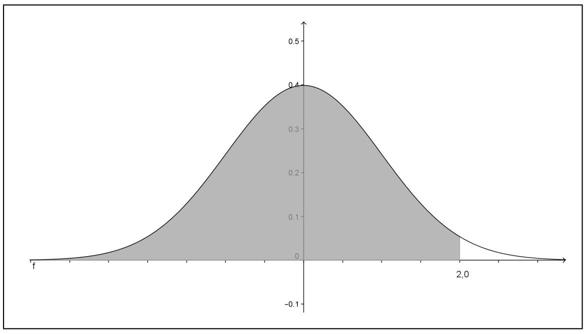


Figura 4.32 – $P(Z \geq -0,5)$

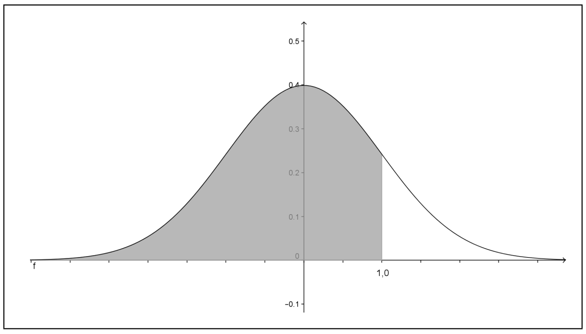


Figura 4.33 – $P(Z \leq 0,5)$

EXEMPLO 4.16 $P(-2,1 \leq Z \leq -1,4)$

Usando os resultados do Exemplo 4.15, temos que

$$P(-2,1 \leq Z \leq -1,4) = P(1,4 \leq Z \leq 2,1) = \Phi(2,1) - \Phi(1,4) = 0,9821 - 0,9192 = 0,0629$$

♦♦

EXEMPLO 4.17 $P(-2,1 \leq Z \leq 1,4)$

Usando os resultados do Exemplo 4.12, temos que

$$\begin{aligned} P(-2,1 \leq Z \leq 1,4) &= \Phi(1,4) - P(Z < -2,1) = \Phi(1,4) - \Phi(-2,1) \\ &= \Phi(1,4) - [1 - \Phi(2,1)] = 0,9192 - [1 - 0,9821] = 0,9013 \end{aligned}$$

♦♦

4.3 A variável aletória $N(\mu; \sigma^2)$

4.3.1 Definição

Seja $Z \sim N(0; 1)$ e vamos definir uma nova variável aleatória $X = g(Z) = \mu + \sigma Z$, em que $\sigma > 0$. Usando o Teorema 3.1, temos que:

$$f_X(x) = f_Z\left(\frac{x-\mu}{\sigma}\right) \times \frac{1}{\sigma} = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \times \frac{1}{\sigma}$$

ou ainda:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

e essa é a densidade normal $N(\mu; \sigma^2)$

DEFINIÇÃO **A densidade** $N(\mu, \sigma^2)$

Uma variável aleatória contínua X, definida para todos os valores em $\mathbb{R}$, tem densidade normal com parâmetros μ e σ^2, onde $-\infty < \mu < \infty$ e $0 < \sigma^2 < \infty$, se sua função densidade de probabilidade é dada por

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \qquad -\infty < x < \infty \quad . \tag{4.13}$$

Usaremos a seguinte notação para indicar que uma variável aleatória X tem distribuição normal com parâmetros μ e σ^2 : $X \sim N(\mu; \sigma^2)$.

4.3.2 Características da curva normal

1. Simétrica em torno de μ; note que $f(\mu - x) = f(\mu + x)$.

2. Assíntotas: $\lim_{x \to -\infty} f(x) = \lim_{x \to \infty} f(x) = 0$; esse resultado segue diretamente do fato de que $\lim_{x\to\infty} e^{-x} = 0$

3. Ponto de máximo

   Para calcular a primeira e segunda derivadas de $f(x)$, devemos lembrar que $(e^x)' = e^x$ e, pela regra da cadeia, $(e^{g(x)})' = e^{g(x)}g'(x)$. Aplicando esses resultados à densidade normal, obtemos que:

$$f'(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]\left[-\frac{1}{2\sigma^2}2(x-\mu)\right] = -f(x)\left(\frac{x-\mu}{\sigma^2}\right) \tag{4.14}$$

   Derivando novamente, obtemos:

$$\begin{aligned} f''(x) &= -f'(x)\left(\frac{x-\mu}{\sigma^2}\right) - f(x)\frac{1}{\sigma^2} = -\left[-f(x)\left(\frac{x-\mu}{\sigma^2}\right)\right]\left[\frac{x-\mu}{\sigma^2}\right] - f(x)\frac{1}{\sigma^2} = \\ &= f(x)\left[\frac{(x-\mu)^2}{\sigma^4}\right] - f(x)\frac{1}{\sigma^2} = f(x)\left[\frac{(x-\mu)^2 - \sigma^2}{\sigma^4}\right] \end{aligned} \tag{4.15}$$

Analisando a equação (4.14) e lembrando que $f(x) > 0$, pode-se ver que:

$$f'(x) = 0 \Leftrightarrow x = \mu$$

e assim, $x = \mu$ é um ponto crítico. Como $f'(x) > 0$ para $x < \mu$ e $f'(x) < 0$ para $x > \mu$, então f é crescente à esquerda de μ e decrescente à direita de μ. Segue, então, que $x = \mu$ é um ponto de máximo e nesse ponto

$$f(\mu) = \frac{1}{\sqrt{2\pi\sigma^2}} \tag{4.16}$$

4. Pontos de inflexão

   Analisando a segunda derivada dada por (4.15), tem-se que:

$$f''(x) = 0 \Leftrightarrow (x-\mu)^2 = \sigma^2 \Leftrightarrow |x-\mu| = \sigma \Leftrightarrow \begin{cases} x = \mu + \sigma \\ x = \mu - \sigma \end{cases} \tag{4.17}$$

   Além disso,

$$\begin{aligned} f''(x) > 0 \quad &\Leftrightarrow \quad (x-\mu)^2 > \sigma^2 \Leftrightarrow |x-\mu| > \sigma \Leftrightarrow \\ &\Leftrightarrow \quad x - \mu > \sigma \quad \text{ou} \quad \mu - x > \sigma \\ &\Leftrightarrow \quad x > \mu + \sigma \quad \text{ou} \quad x < \mu - \sigma \end{aligned} \tag{4.18}$$

   e

$$\begin{aligned} f''(x) \quad &< \quad 0 \Leftrightarrow (x-\mu)^2 < \sigma^2 \Leftrightarrow |x-\mu| < \sigma \Leftrightarrow \\ &\Leftrightarrow \begin{cases} x - \mu < \sigma \\ \mu - x < \sigma \end{cases} \Leftrightarrow \mu - \sigma < x < \mu + \sigma \end{aligned} \tag{4.19}$$

   Logo, $f(x)$ é côncava para cima se $x > \mu + \sigma$ ou $x < \mu - \sigma$ e é côncava para baixo quando $\mu - \sigma < x < \mu + \sigma$.

Na Figura 4.34 é apresentado o gráfico da densidade normal no caso em que $\mu = 3$ e $\sigma^2 = 1$. As linhas pontilhadas passam pelos pontos de inflexão 3 ± 1.

4.3.3 Parâmetros da $N\left(\mu;\sigma^2\right)$

Se $X \sim N\left(\mu;\sigma^2\right)$, então $X = \mu + \sigma Z$, em que $Z \sim N(0;1)$. Das propriedades de média e variância, sabemos que, se X é uma variável aleatória e $k_1 \neq 0$ e k_2 são constantes quaisquer, então

$$\begin{aligned} \mathrm{E}(k_1 X + k_2) &= k_1 \mathrm{E}(X) + k_2 \\ \mathrm{Var}(k_1 X + k_2) &= k_1^2 \mathrm{Var}(X) \end{aligned}$$

Resulta, então, que se $X \sim N\left(\mu;\sigma^2\right)$ então

$$\mathrm{E}(X) = \mu + \sigma\,\mathrm{E}(Z) = \mu + 0 \Rightarrow \mathrm{E}(X) = \mu$$

e

$$\mathrm{Var}(X) = \sigma^2\,\mathrm{Var}(Z) = \sigma^2 \times 1 \Rightarrow \mathrm{Var}(X) = \sigma^2$$

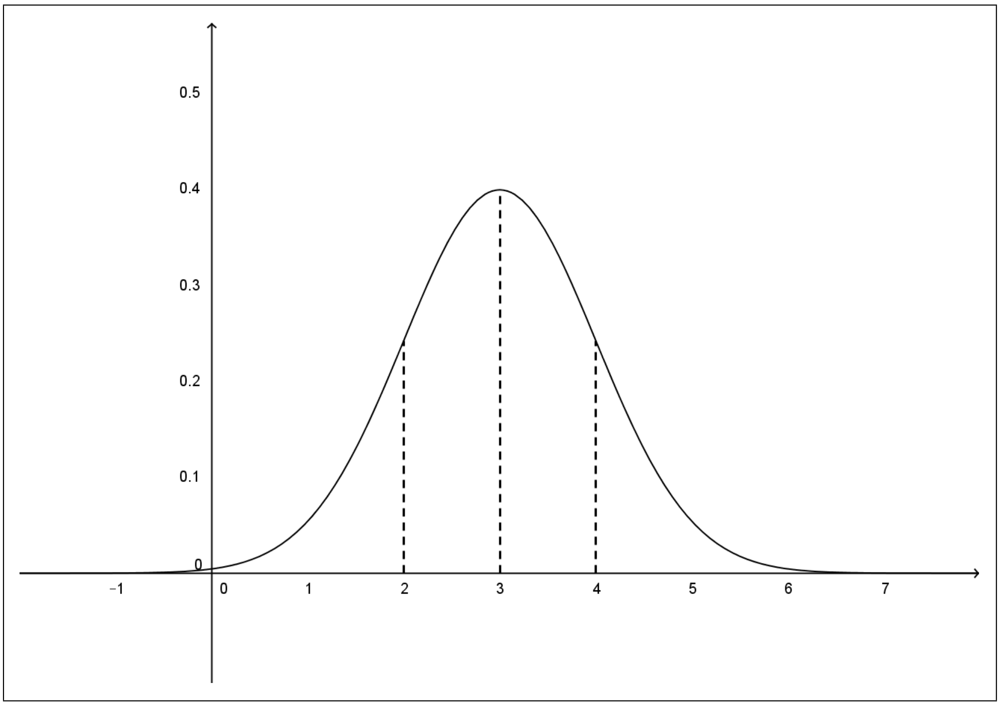


Figura 4.34 – Densidade normal com média $\mu = 3$ e variância $\sigma^2 = 1$

Resumindo:

$$X \sim N\left(\mu; \sigma^2\right) \Longrightarrow \begin{cases} \mathrm{E}(X) = \mu \\ \mathrm{Var}(X) = \sigma^2 \end{cases} \tag{4.20}$$

Os parâmetros da densidade normal são, então, a média e a variância, que são medidas de posição e dispersão, respectivamente. Valores diferentes de μ deslocam o eixo de simetria da curva e valores diferentes de σ^2 mudam a dispersão da curva. Quanto maior σ^2, mais "espalhada" é a curva; mas o ponto de máximo, dado pela equação (4.16), é inversamente proporcional a σ^2. Logo, quanto maior σ^2, mais "espalhada" e mais "achatada" é a curva. A questão é que a forma é sempre a de um "sino". Na Figura 4.35 temos exemplos densidades normais com a mesma variância, mas com médias diferentes. O efeito é o "delocamento" da densidade. Já na Figura 4.36, temos duas densidades com a mesma média, mas variâncias diferentes. O efeito é que a densidade com maior variância é mais dispersa e achatada.

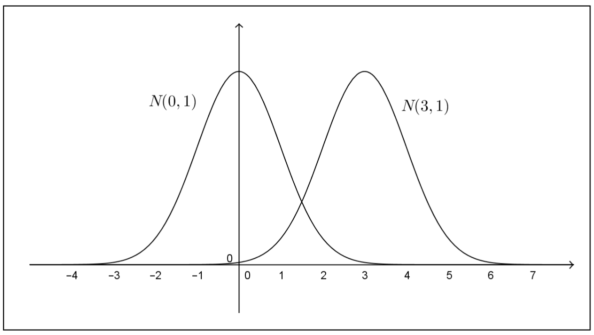


Figura 4.35 – Densidades normais com mesma variância e médias diferentes

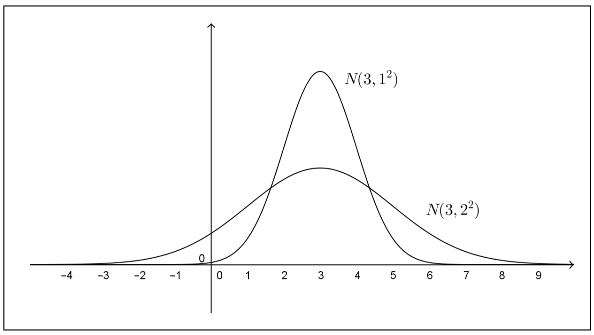


Figura 4.36 – Densidades normais com mesma média e variâncias diferentes

4.3.4 Função de distribuição acumulada

Como no caso da normal padrão, a função de distribuição acumulada não pode ser calculada diretmanete, sendo necessários programas computacionais.Com o auxílio da função DISTR.NORM do Excel foi obtida a **Figura 4.37**. onde temos os gráficos da função de distribuição acumulada para as densidades $N(0,1)$, $N(3,1)$ e $N(3,2)$.

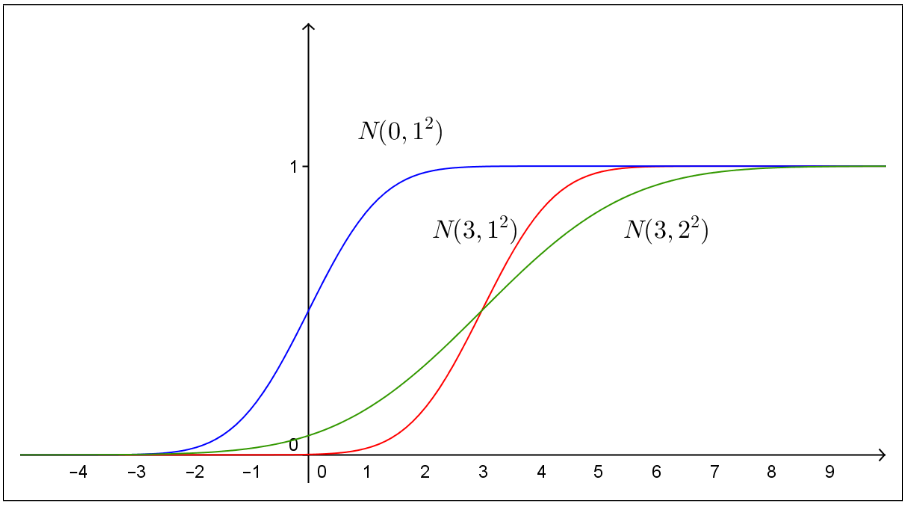


Figura 4.37 – Função de distribuição de várias densidades normais

Note que, pela simetria da densidade em torno da média μ, sempre teremos $\Phi(\mu) = 0,5$.

4.3.5 Cálculos com a distribuição normal

Nesta seção serão apresentados resultados básicos sobre a distribuição normal, que permitirão que você calcule probabilidades associadas a qualquer variável aleatória normal, e isso ampliará o escopo de aplicações práticas.

Na seção anterior, você viu como usar tabelas da distribuição normal padrão para calcular probabilidades associadas à variável $Z \sim N(0;1)$. Essas tabelas, ou softwares especializados, são necessários para fazer os cálculos, pois não existem métodos diretos para calcular áreas sob a curva da densidade normal padrão. Mas as tabelas vistas referiam-se à distribuição $N(0;1)$. Será que teremos que usar uma tabela diferente para outros valores da média μ e do desvio-padrão σ? Felizmente, a resposta é NãO, graças a uma propriedade muito interessante da distribuição normal que estabelece o seguinte resultado:

! Padronização da distribuição normal $N(\mu;\sigma^2)$

Se $X \sim N\left(\mu;\sigma^2\right)$, então

$$Z = \frac{X-\mu}{\sigma} \tag{4.21}$$

tem distribuição $N(0;1)$.

Note que a transformação dada em (4.3.5) é uma transformação linear, que é biunívoca. Vejamos como usar esse resultado para calcular probabilidades de uma v.a. normal qualquer.

Suponhamos, por exemplo, que se deseje calcular $\mathrm{P}(X \leq 3)$, em que $X \sim N(1;2)$, ou seja, X é uma v.a. normal com média 1 e variância 2. Temos a seguinte equivalência de eventos

$$X \leq 3 \quad \Leftrightarrow \quad \frac{X-1}{\sqrt{2}} \leq \frac{3-1}{\sqrt{2}}$$

uma vez que subtraímos a mesma constante e dividimos pela mesma constante positiva em ambos os lados da desigualdade. Mas, pelo resultado acima, $Z = \frac{X-1}{\sqrt{2}} \sim N(0;1)$. Logo,

$$\mathrm{P}(X \leq 3) = \mathrm{P}\left(\frac{X-1}{\sqrt{2}} \leq \frac{3-1}{\sqrt{2}}\right) = \mathrm{P}\left(Z \leq \frac{3-1}{\sqrt{2}}\right)$$

e caímos novamente no cálculo de probabilidades da Normal padrão, que é feito com auxílio das Tabelas 1 e 2, apresentadas na seção anterior. Completando o cálculo, obtemos:

$$\begin{aligned}
\mathrm{P}(X \leq 3) &= \mathrm{P}\left(Z \leq \frac{3-1}{\sqrt{2}}\right) = \mathrm{P}(Z \leq 1,41) \\
&= 0,5 + \mathrm{tab}(1,41) = \Phi(1,41) = 0,9207
\end{aligned}$$

Na Figura 4.38 representa-se a probabilidade $\mathrm{P}(X \leq 3)$ e na Figura 4.39, $\mathrm{P}(Z \leq 1,41)$. Pelo resultado acima, essas duas áreas são iguais.

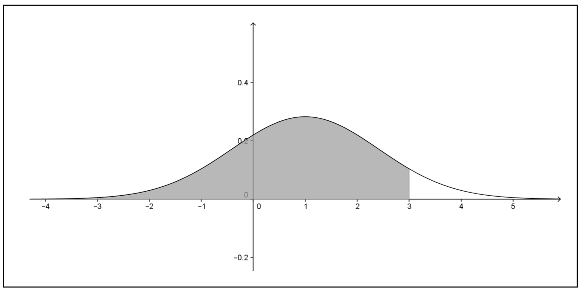


Figura 4.38 – $\mathrm{P}(X \leq 3) - X \sim N(1;2)$

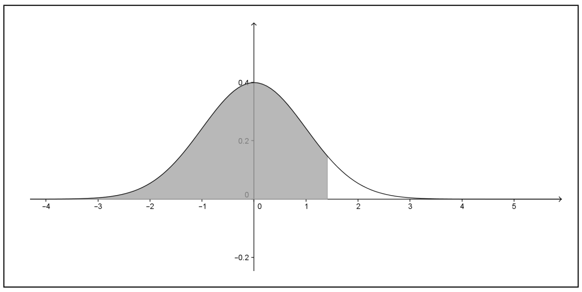


Figura 4.39 – $\mathrm{P}(Z \leq 1,41)$

É interessante lembrar que a transformação dada na equação (4.3.5) corresponde ao cálculo do escore padronizado associado à abcissa x. Assim, cálculos de probabilidades de v.a. normais sempre envolverão o cálculo do escore padronizado da(s) abscissa(s) de interesse.

Agora, vamos apresentar vários exemplos para fixar os conceitos e procedimentos. é importante que você estabeleça a equivalência dos eventos definidos pela distribuição normal de interesse e pela normal padrão. Como antes, faça um esboço do gráfico das curvas normais sombreando a área desejada.

EXEMPLO 4.18 $X \sim N(3;9) - \mathrm{P}(-1 \leq X \leq 4)$

$$\begin{aligned}
\mathrm{P}(-1 \leq X \leq 4) &= \mathrm{P}\left(\frac{-1-3}{\sqrt{9}} \leq \frac{X-3}{\sqrt{9}} \leq \frac{4-3}{\sqrt{9}}\right) \\
&= \mathrm{P}\left(-1,33 \leq Z \leq 0,33\right) \\
&= \Phi(0,33) - \Phi(-1,33) = 0,62930 - 0,09176 \\
&= tab(0,33) + tab(1,33) = 0,12930 + 0,40824 \\
&= 0,53754
\end{aligned}$$

♦♦

EXEMPLO 4.19 $X \sim N(2;5)$ – $P(1 \le X \le 7)$

$$\begin{aligned}
P(1 \le X \le 7) &= P\left(\frac{1-2}{\sqrt{5}} \le \frac{X-2}{\sqrt{5}} \le \frac{7-2}{\sqrt{5}}\right) \\
&= P\left(-0,45 \le Z \le 2,24\right) \\
&= \Phi(2,24) - \Phi(-0,45) = \Phi(2,24) - [1-\Phi(0,45)] = 0,9875 - [1-0,6700] \\
&= tab(2,24) + tab(0,45) = 0,4875 + 0,1700 \\
&= 0,6575
\end{aligned}$$

♦♦

EXEMPLO 4.20 $X \sim N(5,1)$ – $P(X > 7)$

$$\begin{aligned}
P(X > 7) &= P\left(\frac{X-5}{1} > \frac{7-5}{1}\right) \\
&= P(Z > 2) \\
&= 1,0 - \Phi(2,0) = 1,0 - 0,97725 \\
&= 0,5 - tab(2,0) = 0,5 - 0,47725 \\
&= 0,02275
\end{aligned}$$

♦♦

EXEMPLO 4.21 A regra 68-95-99,7

Seja $X \sim N(\mu;\sigma^2)$. Calcule $P(\mu - k\sigma < X < \mu + k\sigma)$, para $k = 1,2,3$.

Solução

Note que essa probabilidade corresponde à probabilidade de X estar a uma distância de k desvios-padrão da média.

$$\begin{aligned}
P(\mu - k\sigma \le X \le \mu + k\sigma) &= P\left(\frac{\mu - k\sigma - \mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{\mu + k\sigma - \mu}{\sigma}\right) \\
&= P(-k \le Z \le k)
\end{aligned}$$

Note que chegamos a uma probabilidade que não depende de μ ou σ, ou seja, esse resultado vale qualquer que seja a distribuição normal.

- $k = 1$

  $$P(\mu - \sigma \le X \le \mu + \sigma) = P(-1 \le Z \le 1) = 2 \times tab(1,0) = 2 \times 0,3414 = 0,6828$$

- $k = 2$

  $$P(\mu - 2\sigma \le X \le \mu + 2\sigma) = P(-2 \le Z \le 2) = 2 \times tab(2,0) = 2 \times 0,4772 = 0,9544$$

- $k = 3$

$$\mathrm{P}(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = \mathrm{P}(-3 \leq Z \leq 3) = 2 \times tab(3,0) = 2 \times 0,4987 = 0,9974$$

Essas probabilidades nos dizem que, para *qualquer* distribuição normal, 68,28% dos valores estão a um desvio-padrão da média, 95,44% estão a dois desvios-padrão e 99,73% dos valores estão a três desvios-padrão da média. Veja a Figura 4.40 para uma ilustração desses resultados.

Lembre-se de que o teorema de Chebyshev fornecia percentuais análogos para qualquer distribuição. Para distribuições normais, os resultados desses três exemplos mostram percentuais mais precisos.

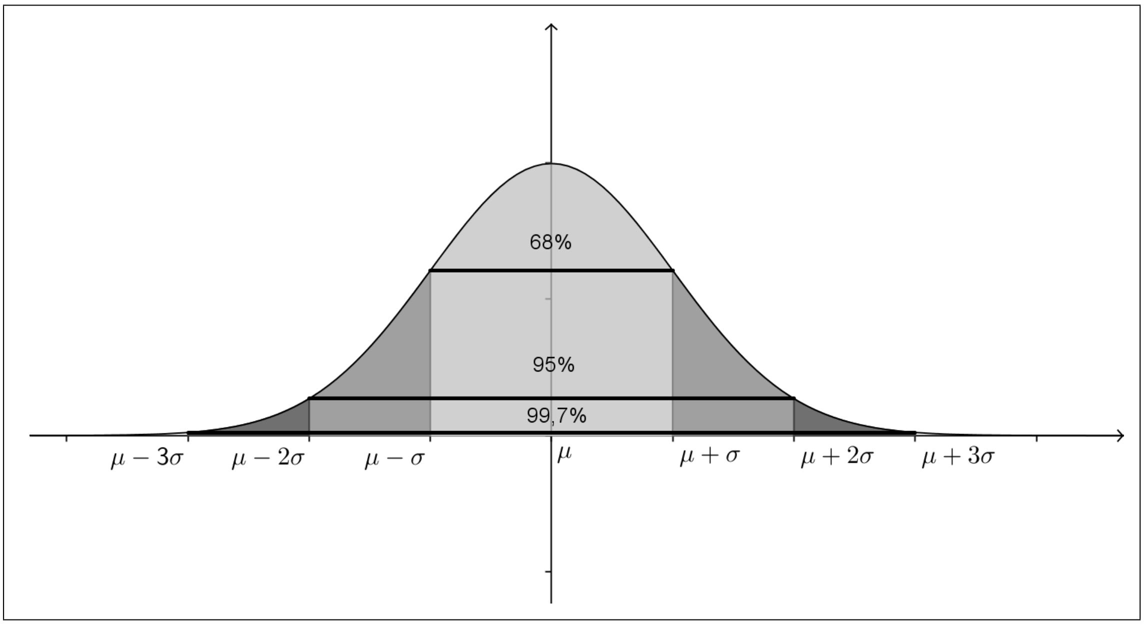


Figura 4.40 – Ilustração da regra 68-95-99,7

4.3.6 Encontrando a abcissa da normal para uma probabilidade específica

Nos exemplos vistos até o momento, consideramos situações em que tínhamos uma abcissa de uma distribuição normal e queríamos a probabilidade associada a essa abcissa. Agora, vamos lidar com a situação inversa: dada uma probabilidade, qual é a abcissa correspondente? Eis algumas situações que envolvem esse tipo de problema:

- Em uma turma de Estatística, os 10% melhores alunos receberão um livro de presente.
- Em uma comunidade, as famílias com as 15% piores rendas irão receber um auxílio da prefeitura.

Como antes, vamos apresentar vários exemplos que ilustram essa situação.

EXEMPLO 4.22 $Z \sim N(0;1)$

Determine o valor de k tal que $\mathrm{P}(Z \leq k) = 0,90$.

Solução

Vamos "traduzir" esse problema: queremos encontrar a abcissa k da normal padrão com $0,90$ de área (probabilidade) à esquerda dela. Como 0,9 é a área à esquerda de k, resulta que k tem que ser maior que zero, isto é, temos que ter $k > 0$. Veja a Figura 4.41: á esquerda de k temos área 0,90 e à esquerda de 0 temos área 0,5. Logo, entre 0 e k temos que ter área 0,40.

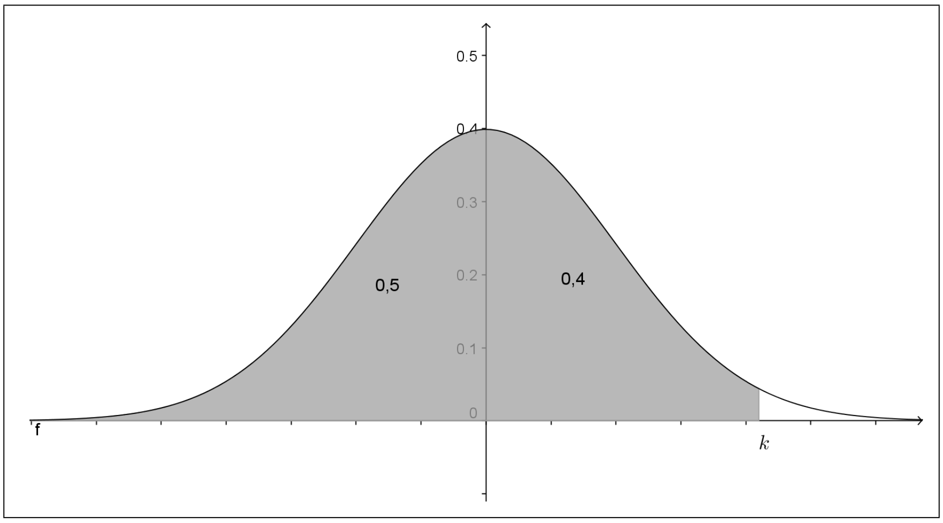


Figura 4.41 – Determinação de k tal que $P(Z \leq k) = 0,90$

Escrevendo essas observações em termos de probabilidade, temos:

$$\begin{aligned}
&P(Z \leq k) = 0,90 \Leftrightarrow \\
&P(Z \leq 0) + P(0 < Z \leq k) = 0,90 \Leftrightarrow \\
&0,5 + P(0 < Z \leq k) = 0,90 \Leftrightarrow \\
&P(0 < Z \leq k) = 0,40 \Leftrightarrow \\
&\mathrm{tab}(k) = 0,40
\end{aligned}$$

Esta última igualdade nos diz que k é a abcissa correspondente ao valor 0,40 na Tabela 1. Para identificar k, temos que buscar no corpo dessa tabela, o valor mais próximo de 0,40. Na linha correspondente ao valor 1,2 encontramos as entradas 0,39973 e 0,40147. Como a primeira está mais próxima de 0,40, olhamos qual é a abcissa correspondente: a linha é 1,2 e a coluna é 8, o que nos dá a abcissa de 1,28, ou seja, $k = 1,28$ e $P(Z \leq 1,28) = 0,90$, completando a solução.

♦♦

Agora vamos olhar o mesmo exemplo, mas para uma distribuição normal qualquer.

EXEMPLO 4.23 $X \sim N(3; 4)$

Determine o valor de k tal que $P(X \leq k) = 0,90$.

Solução

Com a probabilidade à esquerda de k é maior que 0,5, resulta que k tem de ser maior que a média. O primeiro passo na solução é escrever a probabilidade dada em termos da normal padrão.

$$\begin{aligned}
&\mathrm{P}(X \le k) = 0,90 \Leftrightarrow \\
&\mathrm{P}\left(\frac{X-3}{2} \le \frac{k-3}{2}\right) = 0,90 \Leftrightarrow \\
&\mathrm{P}\left(Z \le \frac{k-3}{2}\right) = 0,90 \Leftrightarrow \\
&\mathrm{P}(Z \le 0) + \mathrm{P}\left(0 \le Z \le \frac{k-3}{2}\right) = 0,90 \Leftrightarrow \\
&0,5 + \mathrm{P}\left(0 \le Z \le \frac{k-3}{2}\right) = 0,90 \Leftrightarrow \\
&\mathrm{P}\left(0 \le Z \le \frac{k-3}{2}\right) = 0,40 \Leftrightarrow \\
&tab\left(\frac{k-3}{2}\right) = 0,40 \Leftrightarrow \\
&\frac{k-3}{2} = 1,28 \Leftrightarrow k = 5,56
\end{aligned}$$

EXEMPLO 4.24 $X \sim N(3;4)$

Determine o valor de k tal que $\mathrm{P}(X \le k) = 0,05$.

Solução

à esquerda de k temos 5% da área total; logo, k tem que estar no lado esquerdo, ou seja, temos de ter $k < 3$ e a abcissa padronizada correspondente tem de ser negativa. Vamos escrever a probabilidade dada em termos da normal padronizada:

$$\begin{aligned}
&\mathrm{P}(X \le k) = 0,05 \Leftrightarrow \\
&\mathrm{P}\left(\frac{X-3}{2} \le \frac{k-3}{2}\right) = 0,05 \Leftrightarrow \\
&\mathrm{P}\left(Z \le \frac{k-3}{2}\right) = 0,05
\end{aligned}$$

Como a área (probabilidade) à esquerda de $\frac{k-3}{2}$ é menor que $0,5$, isso significa que $\frac{k-3}{2}$ tem de ser negativo. Veja a Figura 4.42. Para nos adequarmos à tabela disponível, temos de rabalhar com áreas na metade direita da curva densidade, ou seja, temos qde usar a simetria da curva. Veja a Figura 4.43 e note que a abcissa simétrica a $\dfrac{k-3}{2}$ é $-\dfrac{k-3}{2} = \dfrac{3-k}{2}$.

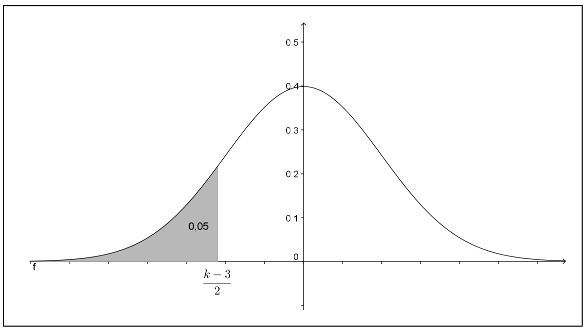


Figura 4.42 – $\mathrm{P}(Z \le \frac{k-3}{2}) = 0,05$

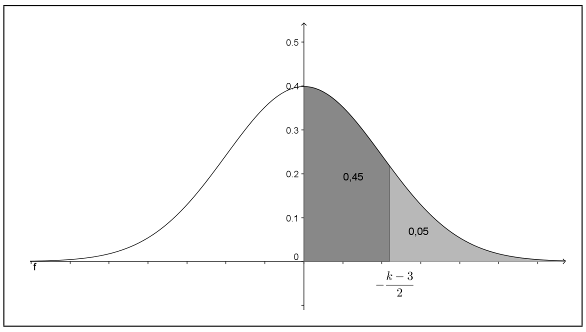


Figura 4.43 – $\mathrm{P}(Z \ge -\frac{k-3}{2}) = 0,05$

Temos, então, as seguintes probabilidades equivalentes:

$$\begin{aligned}
\mathrm{P}\left(Z \leq \frac{k-3}{2}\right) &= 0,05 \Leftrightarrow \\
\mathrm{P}\left(Z \geq -\frac{k-3}{2}\right) &= 0,05 \Leftrightarrow \\
\mathrm{P}\left(0 \leq Z \leq -\frac{k-3}{2}\right) &= 0,45 \Leftrightarrow \\
\mathrm{tab}\left(-\frac{k-3}{2}\right) &= 0,45
\end{aligned}$$

O valor mais próximo de 0,45 no corpo da Tabela 1 é 0,4495 que corresponde à abcissa 1,64, e isso nos dá que

$$-\frac{k-3}{2} = 1,64 \Rightarrow k = -0,28$$

EXEMPLO 4.25 $X \sim N(3; 4)$

Deermine o valor de k tal que $\mathrm{P}(|\,X - 3\,| \leq k) = 0,95$.

Solução

Vamos relembrar as propriedades da função módulo. Para isso, veja a Figura 4.44. As duas retas que definem a função $f(x) = |x|$ são $y = x$ e $y = -x$. Os segmentos traçados no gráfico mostram que $|x| = k \Leftrightarrow x = k$ ou $x = -k$. Os valores de y abaixo do segmento horizontal correspondem a valores de $y = |x| \leq k$ e esses valores de y estão associados a valores x tais que $-k \leq x \leq k$. De forma análoga, valores de y acima do segmento horizontal correspondem a valores de $y = |x| > k$ e esses valores de y estão associados a valores x tais que $x > k$ ou $x < -k$. Resumindo:

$$|x| = k \Leftrightarrow \begin{cases} \text{x=k} \\ \text{ou} \\ \text{x=-k} \end{cases} \tag{4.22}$$

$$|x| < k \Leftrightarrow -k < x < k \tag{4.23}$$

$$|x| > k \Leftrightarrow \begin{cases} \text{x>k} \\ \text{ou} \\ \text{x<-k} \end{cases} \tag{4.24}$$

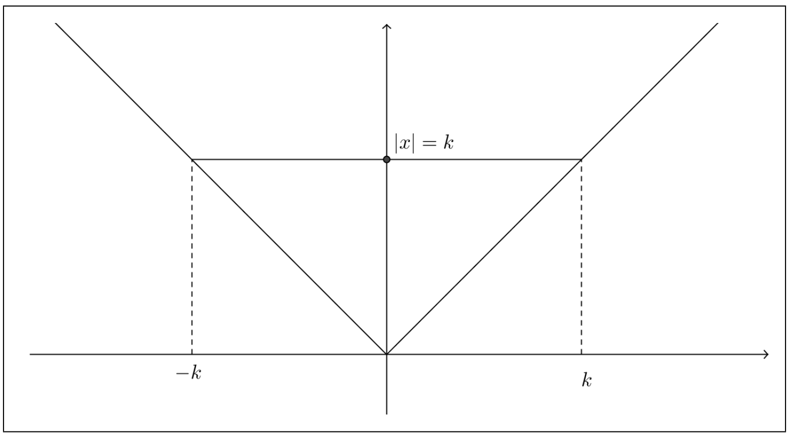


Figura 4.44 – Determinação de k tal que $\mathrm{P}(Z \leq k) = 0,90$

Agora, vamos usar esses resultados para resolver o exemplo.

$$
\begin{aligned}
&\mathrm{P}(|\,X-3\,| \le k) = 0,95 \Leftrightarrow \\
&\mathrm{P}(-k \le X-3 \le k) = 0,95 \Leftrightarrow \\
&\mathrm{P}\,(3-k \le X \le k+3) = 0,95 \Leftrightarrow \\
&\mathrm{P}\left(\frac{3-k-3}{2} \le \frac{X-3}{2} \le \frac{k+3-3}{2}\right) = 0,95 \Leftrightarrow \\
&\mathrm{P}\left(-\frac{k}{2} \le Z \le \frac{k}{2}\right) = 0,95
\end{aligned}
$$

Veja a Figura 4.45 para entender que

$$
\begin{aligned}
&\mathrm{P}\left(-\frac{k}{2} \le Z \le \frac{k}{2}\right) = 0,95 \Leftrightarrow \\
&\mathrm{P}\left(-\frac{k}{2} \le Z \le 0\right) + \mathrm{P}\left(0 \le Z \le \frac{k}{2}\right) = 0,95 \Leftrightarrow \\
&2 \times \mathrm{P}\left(0 \le Z \le \frac{k}{2}\right) = 0,95 \Leftrightarrow \\
&\mathrm{P}\left(0 \le Z \le \frac{k}{2}\right) = 0,475 \Leftrightarrow \\
&\mathrm{tab}\left(\frac{k}{2}\right) = 0,475 \Leftrightarrow \\
&\frac{k}{2} = 1,96 \Leftrightarrow k = 3,92
\end{aligned}
$$

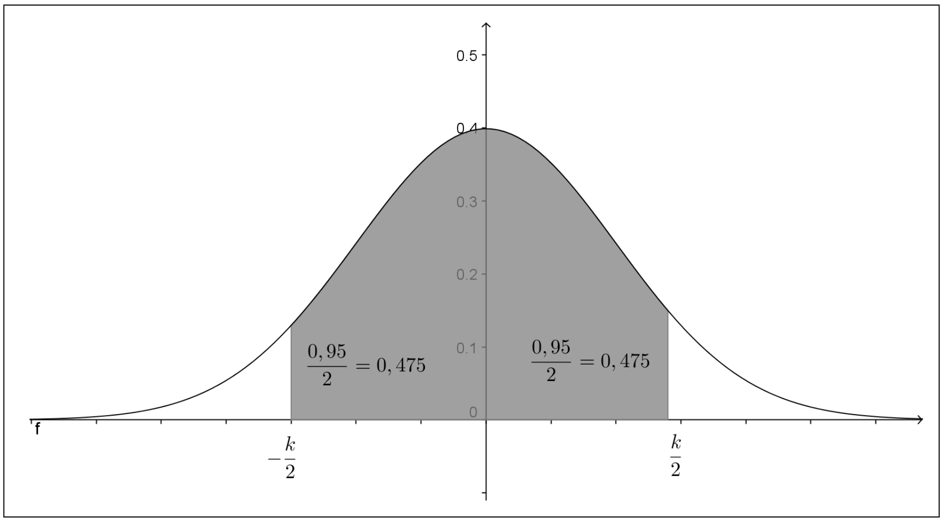


Figura 4.45 – Determinação de k tal que $\mathrm{P}(|ZZ \le \frac{k}{2}) = 0,95$

♦♦

4.3.7 Exemplos de Aplicação da Distribuição Normal

A distribuição normal é um modelo probabilístico que se aplica a diversas situações práticas. Vamos finalizar este capítulo com alguns exemplos práticos, mas, na última parte do curso, você verá mais aplicações no contexto da inferência estatística, em que decisões têm de ser tomadas com base nos resultados obtidos a partir de uma amostra.

EXEMPLO 4.26 Conta de celular

O consumo mensal em minutos por conta de celular em uma região é uma variável aleatória normal com média 36 e desvio padrão 12.

(a) Qual é a probabilidade de uma pessoa desta região usar o telefone celular por menos de 48 minutos?

(b) Qual é a probabilidade de uma pessoa desta região usar o telefone celular por mais de 30 minutos?

(c) Qual o tempo mínimo que alguém deve gastar ao telefone no mês para estar entre os 5% que MAIS usam o celular?

(d) Qual o tempo máximo que alguém deve gastar ao telefone no mês para estar entre os 12% que MENOS usam o celular?

Solução

X = consumo em minutos; $X \sim N(36; 12^2)$

1. $P(X < 48) = P\left(Z < \frac{48-36}{12}\right) = P(Z < 1) = 0,5 + tab(1,0) = \Phi(1,0) = 0,84134$

2. $P(X > 30) = P\left(Z > \frac{30-36}{12}\right) = P(Z > -0,5) = 0,5 + tab(0,5) = \Phi(0,5) = 0,69146$

3. Seja m o tempo mínimo.

$$\begin{aligned} P(X \geq m) = 0,05 \quad &\Leftrightarrow \quad P\left(Z > \frac{m-36}{12}\right) = 0,05 \Leftrightarrow tab\left(\frac{m-36}{12}\right) = 0,45 \\ &\Leftrightarrow \quad \frac{m-36}{12} = 1,64 \Leftrightarrow m = 55,68 \end{aligned}$$

A pessoa tem que falar pelo menos 55,68 minutos para estar entre os 5% que mais usam o celular.

4. Seja M o tempo máximo.

$$\begin{aligned} P(X \leq M) = 0,12 \quad &\Leftrightarrow \quad P\left(Z \leq \frac{M-36}{12}\right) = 0,12 \Leftrightarrow P\left(Z \geq \frac{36-M}{12}\right) = 0,12 \\ &\Leftrightarrow \quad tab\left(\frac{36-M}{12}\right) = 0,5 - 0,12 \Leftrightarrow \frac{36-M}{12} = 1,175 \Leftrightarrow M = 21,9 \end{aligned}$$

A pessoa tem que falar no máximo 21,9 minutos para estar entre os 12% que menos usam o celular.

EXEMPLO 4.27 Saldo bancário

O saldo médio dos clientes de um banco é uma v.a. normal com média R$ 2.000, 00 e desvio-padrão R$ 250,00. Os clientes com os 10% maiores saldos médios recebem tratamento VIP, enquanto aqueles com os 5% menores saldos médios receberão propaganda extra para estimular maior movimentação da conta.

(a) Quanto você precisa de saldo médio para se tornar um cliente VIP?

(b) Abaixo de qual saldo médio o cliente receberá a propaganda extra?

Solução

Seja X = "saldo médio"; é dado que $X \sim N(2000; 250^2)$.

(a) Temos que determinar o valor de k tal que $P(X \geq k) = 0,10$. Note que isso equivale a calcular o 90º percentil da distribuição. A área à esquerda de k tem de ser 0,90; logo, k tem de ser maior que a média.

$$
\begin{aligned}
P(X \geq k) &= 0,10 \Leftrightarrow \\
P\left(\frac{X-2000}{250} \geq \frac{k-2000}{250}\right) &= 0,10 \Leftrightarrow \\
P\left(\frac{X-2000}{250} \leq \frac{k-2000}{250}\right) &= 0,90 \Leftrightarrow \\
P\left(Z \leq \frac{k-2000}{250}\right) &= 0,90 \Leftrightarrow \\
P(Z \leq 0) + P\left(0 \leq Z \leq \frac{k-2000}{250}\right) &= 0,90 \Leftrightarrow \\
P\left(0 \leq Z \leq \frac{k-2000}{250}\right) &= 0,90 - 0,50 \Leftrightarrow \\
\text{tab}\left(\frac{k-2000}{250}\right) &= 0,40 \Leftrightarrow \\
\frac{k-2000}{250} &= 1,28 \Leftrightarrow k = 2320
\end{aligned}
$$

Os clientes com saldo médio maior ou igual a R\$ 2.320,00 terão tratamento VIP.

(b) Temos de determinar o valor de k tal que $P(X \leq k) = 0,05$. Note que isso equivale a calcular o 5º percentil da distribuição. A área à esquerda de k tem de ser 0,05; logo, k tem de ser menor que a média. Na solução, teremos que usar a simetria da distribuição, invertendo o sinal da abcissa, para lidarmos com área na metade direita da função densidade.

$$
\begin{aligned}
P(X \leq k) &= 0,05 \Leftrightarrow \\
P\left(\frac{X-2000}{250} \leq \frac{k-2000}{250}\right) &= 0,05 \Leftrightarrow \\
P\left(Z \geq -\frac{k-2000}{250}\right) &= 0,05 \Leftrightarrow \\
P\left(Z \geq \frac{2000-k}{250}\right) &= 0,05 \Leftrightarrow \\
\text{tab}\left(\frac{2000-k}{250}\right) &= 0,45 \Leftrightarrow \\
\frac{2000-k}{250} &= 1,64 \Leftrightarrow k = 1590
\end{aligned}
$$

Os clientes com saldo médio inferior a R\$ 1.590,00 receberão a propaganda extra.

Na Figura 4.46 ilustra-se a solução do exercício.

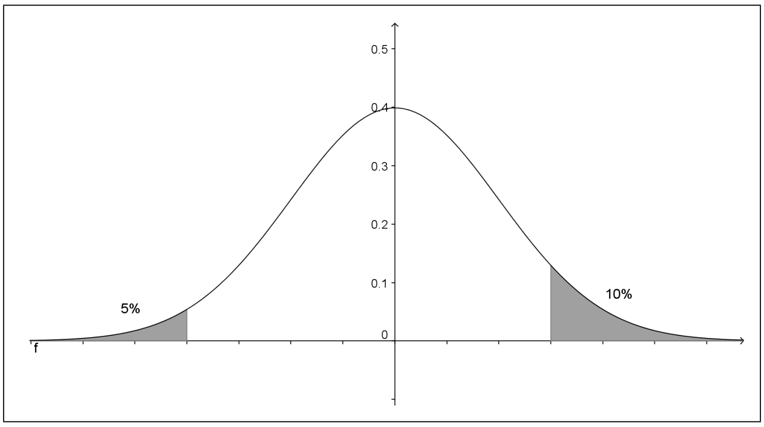


Figura 4.46 – Solução do Exemplo 4.27

♦♦

EXEMPLO 4.28 Regulagem de máquinas

Uma máquina de empacotar determinado produto oferece variações de peso que se distribuem segundo uma distribuição normal com desvio padrão de 20 gramas. Em quanto deve ser regulado o peso médio desses pacotes para que apenas 10% deles tenham menos que 500 gramas?

Solução

Esse é um exemplo clássico de aplicação da distribuição normal. Seja X o peso dos pacotes em gramas. Então, $X \sim N(\mu; 400)$. Temos de ter $\mathrm{P}(X \leq 500) = 0,10$. Note que o peso médio tem de ser superior a 500 g. Note, na solução, a inversão do sinal da abcissa!

$$
\begin{aligned}
\mathrm{P}(X \leq 500) &= 0,10 \Leftrightarrow \\
\mathrm{P}\left(\frac{X-\mu}{20} \leq \frac{500-\mu}{20}\right) &= 0,10 \Leftrightarrow \\
\mathrm{P}\left(Z \leq \frac{500-\mu}{20}\right) &= 0,10 \Leftrightarrow \\
\mathrm{P}\left(Z \geq -\frac{500-\mu}{20}\right) &= 0,10 \Leftrightarrow \\
\mathrm{P}\left(Z \geq \frac{\mu-500}{20}\right) &= 0,10 \Leftrightarrow \\
\mathrm{tab}\left(\frac{\mu-500}{20}\right) &= 0,40 \Leftrightarrow \\
\frac{\mu-500}{20} &= 1,28 \Leftrightarrow \mu = 525,6
\end{aligned}
$$

A máquina tem de ser regulada com um peso médio de 525,6g para que apenas 10% dos pacotes tenham peso inferior a 500g. Veja a Figura 4.47.

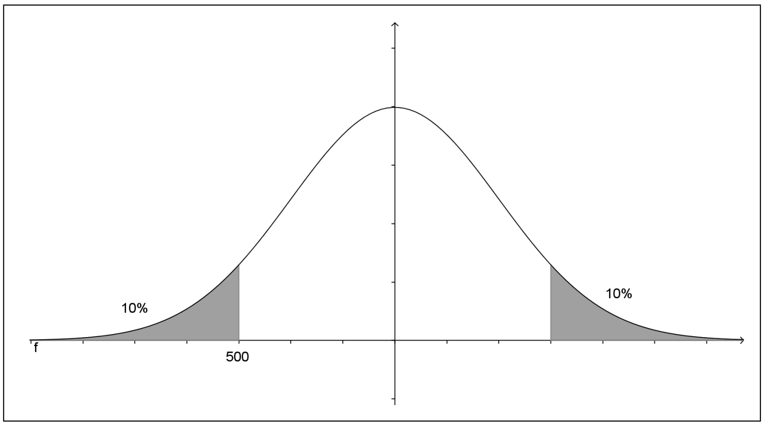


Figura 4.47 – Solução do Exemplo 4.28

♦♦

EXEMPLO 4.29 Mais sobre regulagem de máquinas

Uma máquina fabrica tubos metálicos cujos diâmetros podem ser considerados uma variável aleatória normal com média 200mm e desvio-padrão 2mm. Verifica-se que 15% dos tubos estão sendo rejeitados como grandes e 10% como pequenos.

(a) Quais são as tolerâncias de especificação para esse diâmetro?

(b) Mantidas essas especificações, qual deverá ser a regulagem média da máquina para que a rejeição por diâmetro grande seja nula? Nesse caso, qual será a porcentagem de rejeição por diâmetro pequeno?

Solução

Seja $D =$ diâmetro dos tubos. Então $D \sim N(200, 2^2)$.

(a) Sejam k_I e k_S as especificações inferior e superior, respectivamente. Isso significa que tubos com diâmetro menor que k_I são rejeitados como pequenos e tubos com diâmetro maior que k_S são rejeitados como grandes.

$$
\begin{aligned}
\mathrm{P}(D < k_I) = 0,10 &\Rightarrow \\
\mathrm{P}\left(\frac{D-200}{2} < \frac{k_I - 200}{2}\right) = 0,10 &\Rightarrow \\
\mathrm{P}\left(Z < \frac{k_I - 200}{2}\right) = 0,10 &\Rightarrow \\
\mathrm{P}\left(Z > -\frac{k_I - 200}{2}\right) = 0,10 &\Rightarrow \\
\mathrm{P}\left(Z > \frac{200 - k_I}{2}\right) = 0,10 &\Rightarrow \\
\mathrm{P}\left(0 \leq Z < \frac{200 - k_I}{2}\right) = 0,40 &\Rightarrow \\
\mathrm{tab}\left(\frac{200 - k_I}{2}\right) = 0,40 &\Rightarrow \\
\frac{200 - k_I}{2} = 1,28 \Rightarrow k_I = 197,44 &
\end{aligned}
$$

$$\begin{aligned}
\mathrm{P}(D > k_S) = 0,15 &\Rightarrow \\
\mathrm{P}\left(\frac{D-200}{2} > \frac{k_S-200}{2}\right) = 0,15 &\Rightarrow \\
\mathrm{P}\left(0 \le Z < \frac{k_S-200}{2}\right) = 0,35 &\Rightarrow \\
\mathrm{tab}\left(\frac{k_S-200}{2}\right) = 0,35 &\Rightarrow \\
\frac{k_S-200}{2} = 1,03 \Rightarrow k_S &= 202,06
\end{aligned}$$

Logo, tubos com diâmetro menor que 197,44 cm são rejeitados como pequenos e tubos com diâmetros maiores que 202,06 cm são rejeitados como grandes.

(b) Com a nova regulagem, temos que $D \sim N(\mu; 2^2)$ e μ deve ser tal que

$$\begin{aligned}
\mathrm{P}(D > 202,06) = 0 &\Rightarrow \\
\mathrm{P}\left(\frac{D-\mu}{2} > \frac{202,06-\mu}{2}\right) = 0 &\Rightarrow \\
\mathrm{P}\left(Z > \frac{202,06-\mu}{2}\right) = 0 &\Rightarrow \\
\mathrm{P}\left(0 \le Z \le \frac{202,06-\mu}{2}\right) = 0,5 &\Rightarrow \\
\mathrm{tab}\left(\frac{202,06-\mu}{2}\right) = 0,5 &\Rightarrow \\
\frac{202,06-\mu}{2} \simeq 4,5 \Rightarrow \mu &\simeq 193,06
\end{aligned}$$

Com essa média, a porcentagem de rejeição por diâmetro pequeno é

$$\begin{aligned}
\mathrm{P}(D < 197,44) &= \mathrm{P}\left(\frac{D-193,06}{2} < \frac{197,44-193,06}{2}\right) \\
&= \mathrm{P}(Z < 2,19) \\
&= \mathrm{P}(Z \le 0) + \mathrm{P}(0 < Z < 2,19) \\
&= 0,5 + tab(2,19) = 0,9857
\end{aligned}$$

Com essa nova regulagem, a rejeição por diâmetro grande é nula, mas a rejeição por diâmetro pequeno é muito alta! Veja as Figuras 4.48 e 4.49, nas quais ficam claros os resultados obtidos.

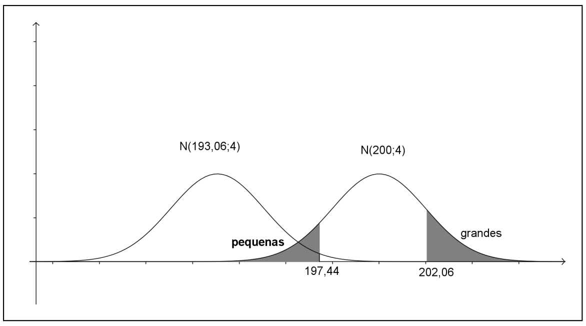


Figura 4.48 – Exemplo 4.29 - tubos pequenos e grandes na regulagem original

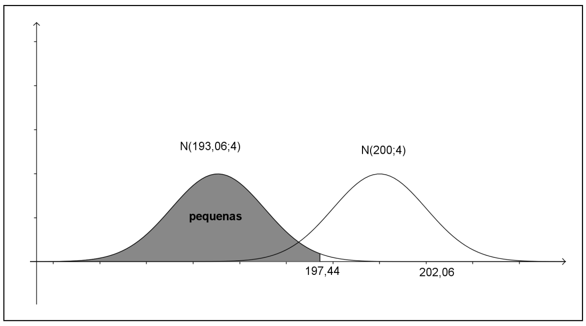


Figura 4.49 – Exemplo 4.29 - regulagem com 0% de tubos grandes

♦♦

EXEMPLO 4.30 Troca de lâmpadas

Em um grande complexo industrial, o departamento de manutenção tem instruções para substituir as lâmpadas antes que se queimem. Os registros indicam que a duração das lâmpadas, em horas, tem distribuição normal, com média de 900 horas e desvio-padrão de 75 horas. Quando devem ser trocadas as lâmpadas, de modo que no máximo 5% delas queimem antes de serem trocadas?

Solução

Seja $T =$ "tempo de duração (em horas) das lâmpadas"; então, $T \sim N(900; 75^2)$. Temos que determinar t tal que $\mathrm{P}(T \leq t) = 0,05$.

$$\begin{aligned}
\mathrm{P}(T \leq t) &= 0,05 \Leftrightarrow \\
\mathrm{P}\left(\frac{T-900}{75} \leq \frac{t-900}{75}\right) &= 0,05 \Leftrightarrow \\
\mathrm{P}\left(Z \geq -\frac{t-900}{75}\right) &= 0,05 \Leftrightarrow \\
\mathrm{P}\left(Z \geq \frac{900-t}{75}\right) &= 0,05 \Leftrightarrow \\
\mathrm{P}\left(0 \leq Z \leq \frac{900-t}{75}\right) &= 0,45 \Leftrightarrow \\
tab\left(\frac{900-t}{75}\right) &= 0,45 \Leftrightarrow \\
\frac{900-t}{75} &= 1,64 \Leftrightarrow t = 777
\end{aligned}$$

As lâmpadas devem ser trocadas com 777 horas de uso para que apenas 5% se queimem antes da troca.

Aqui cabe a seguinte observação: em geral, não é apropriado utilizar-sea distribuição normal para modelar o tempo de sobrevivência de lâmpadas ou equipamentos em geral. Modelos tipo exponencial são mais adequados, pois atribuem probabilidade alta de sobrevivência no início da vida do equipamento e probabilidade decrescente à medida que o equipamento envelhece.

EXEMPLO 4.31 Regulagem de máquinas – controle da dispersão

Uma enchedora automática enche garrafas de acordo com uma distribuição normal de média 1.000 ml. Deseja-se que no máximo uma garrafa em cada 100 saia com menos de 990ml. Qual é o maior desvio padrão tolerável?

Solução

Seja $X =$ "conteúdo da garrafa (em ml)". Então $X \sim N(1000; \sigma^2)$ e queremos determinar o maior σ que forneça $\mathrm{P}(X < 990) \leq 0,01$.

Na Figura 4.50 ilustra-se a situação considerada no problema, em que k corresponde ao valor de corte (990 no problema) e a área sombreada corresponde à probabilidade máxima desejada α (0,01 no problema). A curva com linha preta sólida corresponde à distribuição $N(1000;\sigma_0^2)$ tal que $P(X < k) = \alpha$. A curva tracejada em azul representa distribuições normais com desvio padrão menor que σ_0; para essas distribuições, a área abaixo de k é menor que α, ou seja, se $\sigma < \sigma_0$, a distribuição de X é tal que $P(X < k) < \alpha$. Já a curva pontilhada em vermelho representa distribuições normais com desvio padrão maior que σ_0 e para todas essas distribuições, a área abaixo de k é maior que α. Logo, o valor máximo de σ é aquele correspondente à distribuição $N(1000;\sigma^2)$ tal que a área abaixo de k é exatamente igual a α.

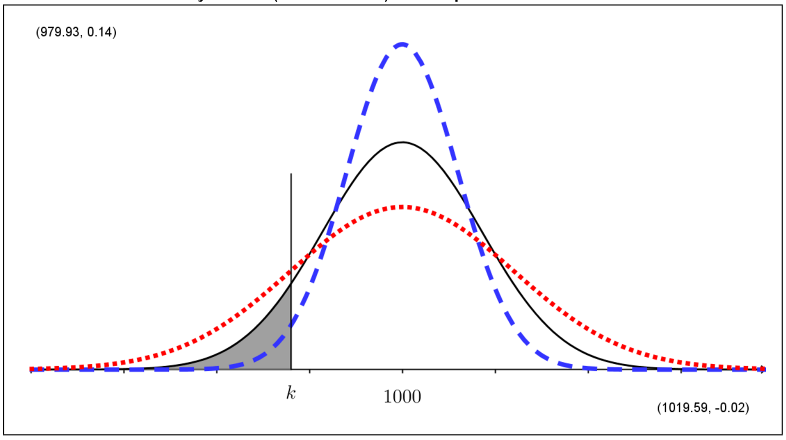


Figura 4.50 – Solução do Exemplo 4.31

Em termos do nosso problema, temos que determinar o desvio padrão σ da distribuição normal com média 1000 que deixa probabilidade 0,01 abaixo de 990, ou seja:

$$
\begin{aligned}
X \sim N(1000;\sigma^2) \text{ tal que } P(X < 990) \le 0,01 &\Leftrightarrow \\
P\left(Z < \frac{990-1000}{\sigma}\right) \le 0,01 &\Leftrightarrow \\
P\left(Z > -\frac{990-1000}{\sigma}\right) \le 0,01 &\Leftrightarrow \\
P\left(Z > \frac{10}{\sigma}\right) \le 0,01 &\Leftrightarrow \\
0,5 - tab\left(\frac{10}{\sigma}\right) \le 0,01 &\Leftrightarrow \\
tab\left(\frac{10}{\sigma}\right) \ge 0,49 &\Leftrightarrow \\
\frac{10}{\sigma} \ge 2,33 \Leftrightarrow \sigma \le \frac{10}{2,33} &= 4,2918
\end{aligned}
$$

♦♦

4.4 A distribuição log-normal

4.4.1 Definição

DEFINIÇÃO Distribuição log-normal

Seja $X \sim N(\mu, \sigma^2)$. Se $Y = e^X$ então a variável aleatória Y tem distribuição log-normal com parâmetros μ e σ^2. Reciprocamente, se Y tem distribuição log-normal, então $X = \ln Y$ tem distribuição $N(\mu, \sigma^2)$.

Vamos calcular a função densidade de probabilidade de uma variável aleatória log-normal a partir de sua função de distribuição acumulada. Note que Y só pode assumir valores positivos. Temos que:

$$\begin{aligned} F_Y(y) &= \Pr(Y \leq y) = \Pr\left(e^X \leq y\right) = \Pr(X \leq \ln y) = \\ &= \Pr\left(\frac{X-\mu}{\sigma} \leq \frac{\ln y - \mu}{\sigma}\right) = \Pr\left(Z \leq \frac{\ln y - \mu}{\sigma}\right) \\ &= \Phi\left(\frac{\ln y - \mu}{\sigma}\right) \qquad y > 0 \end{aligned}$$

Sabemos que $f_Y(y) = F'(y)$ e, também, no caso da normal padrão, $\Phi'(z) = \varphi(z)$. Logo, pela regra da cadeia,

$$\begin{aligned} f_Y(y) &= \Phi'\left(\frac{\ln y - \mu}{\sigma}\right) \times \frac{1}{\sigma y} = \varphi\left(\frac{\ln y - \mu}{\sigma}\right) \times \frac{1}{\sigma y} = \\ &= \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2\right] \times \frac{1}{\sigma y} \end{aligned}$$

ou ainda:

$$f_Y(y) = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2\right] \qquad y > 0$$

4.4.2 Esperança

A esperança de Y é:

$$E(Y) = \int_0^\infty y \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{\ln y - \mu}{\sigma}\right)^2\right] dy$$

Fazendo a mudança de variável

- $t = \ln y$

  temos que

- $y = 0 \Rightarrow t = -\infty$

- $y = \infty \Rightarrow t = \infty$

- $dt = \frac{1}{y} dy$

- $y = e^t$

e, portanto

$$\begin{aligned}
\mathrm{E}(Y) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^t \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right] dt = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2 + t\right] dt \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{t^2 - 2t\mu + \mu^2 - 2\sigma^2 t}{2\sigma^2}\right] dt = \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{t^2 - 2t\left(\mu + \sigma^2\right) + \mu^2}{2\sigma^2}\right] dt \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{t^2 - 2t\left(\mu + \sigma^2\right)}{2\sigma^2}\right] \exp\left(-\frac{\mu^2}{2\sigma^2}\right) dt = \\
&= \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{t^2 - 2t\left(\mu + \sigma^2\right) + \left(\mu + \sigma^2\right)^2 - \left(\mu + \sigma^2\right)^2}{2\sigma^2}\right] dt \\
&= \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left\{-\frac{\left[t - \left(\mu + \sigma^2\right)\right]^2}{2\sigma^2}\right\} \exp\left(\frac{\left(\mu + \sigma^2\right)^2}{2\sigma^2}\right) dt \\
&= \exp\left(-\frac{\mu^2}{2\sigma^2} + \frac{\left(\mu + \sigma^2\right)^2}{2\sigma^2}\right) \left[\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left\{-\frac{\left[t - \left(\mu + \sigma^2\right)\right]^2}{2\sigma^2}\right\} dt\right]
\end{aligned}$$

Mas o termo entre os colchetes externos é a integral de uma densidade normal com média $\lambda = \left(\mu + \sigma^2\right)$ e variância σ^2; logo, essa integral é 1 e, portanto:

$$\mathrm{E}(Y) = \exp\left(-\frac{\mu^2}{2\sigma^2} + \frac{\left(\mu + \sigma^2\right)^2}{2\sigma^2}\right) = \exp\left(\frac{-\mu^2 + \mu^2 + 2\mu\sigma^2 + \sigma^4}{2\sigma^2}\right) \Rightarrow$$

$$\mathrm{E}(Y) = \exp\left(\mu + \frac{\sigma^2}{2}\right) \tag{4.25}$$

4.4.3 Variância

Vamos calcular de modo análogo $E(Y^2)$, usando a mesma transformação:

$$\begin{aligned}
\mathrm{E}(Y^2) &= \int_0^\infty y^2 \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{\ln y-\mu}{\sigma}\right)^2\right] dy \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty e^{2t} \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right] dt = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2 + 2t\right] dt \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left[-\frac{t^2-2t\mu+\mu^2-4\sigma^2 t}{2\sigma^2}\right] dt = \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left[-\frac{t^2-2t\left(\mu+2\sigma^2\right)+\mu^2}{2\sigma^2}\right] dt \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left[-\frac{t^2-2t\left(\mu+2\sigma^2\right)}{2\sigma^2}\right] \exp\left(-\frac{\mu^2}{2\sigma^2}\right) dt = \\
&= \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left[-\frac{t^2-2t\left(\mu+2\sigma^2\right)+\left(\mu+2\sigma^2\right)^2-\left(\mu+2\sigma^2\right)^2}{2\sigma^2}\right] dt \\
&= \exp\left(-\frac{\mu^2}{2\sigma^2}\right) \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left\{-\frac{\left[t-\left(\mu+2\sigma^2\right)\right]^2}{2\sigma^2}\right\} \exp\left(\frac{\left(\mu+2\sigma^2\right)^2}{2\sigma^2}\right) dt \\
&= \exp\left(-\frac{\mu^2}{2\sigma^2}+\frac{\left(\mu+2\sigma^2\right)^2}{2\sigma^2}\right) \left[\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^\infty \exp\left\{-\frac{\left[t-\left(\mu+2\sigma^2\right)\right]^2}{2\sigma^2}\right\} dt\right]
\end{aligned}$$

Como antes, o termo entre os colchetes externos é 1 porque é a integral de uma densidade normal com média $\mu+2\sigma^2$ e variância σ^2. Logo,

$$\begin{aligned}
\mathrm{E}(Y^2) &= \exp\left(-\frac{\mu^2}{2\sigma^2}+\frac{\left(\mu+2\sigma^2\right)^2}{2\sigma^2}\right) = \exp\left(\frac{-\mu^2+\mu^2+4\mu\sigma^2+4\sigma^4}{2\sigma^2}\right) \Rightarrow \\
& E(Y^2) = \exp\left(2\mu+2\sigma^2\right)
\end{aligned}$$

e

$$\begin{aligned}
\mathrm{Var}(Y) &= \exp\left(2\mu+2\sigma^2\right) - \left[\exp\left(\mu+\frac{\sigma^2}{2}\right)\right]^2 = \\
&= \exp\left(2\mu+2\sigma^2\right) - \exp\left[2\left(\mu+\frac{\sigma^2}{2}\right)\right] = \\
&= \exp\left(2\mu+2\sigma^2\right) - \exp\left(2\mu+\sigma^2\right) \\
&= \exp\left(2\mu+\sigma^2\right)\left[\frac{\exp\left(2\mu+2\sigma^2\right)}{\exp\left(2\mu+\sigma^2\right)}-1\right] = \\
&= \exp\left(2\mu+\sigma^2\right)\left[\exp\left(2\mu+2\sigma^2-2\mu-\sigma^2\right)-1\right] = \\
&= \exp\left(2\mu+\sigma^2\right)\left[\exp\sigma^2-1\right]
\end{aligned}$$

Definindo $m = \mathrm{E}(X) = \exp\left(\mu+\frac{\sigma^2}{2}\right)$, temos que $m^2 = \exp\left(2\mu+\sigma^2\right)$. Logo,

$$\mathrm{Var}(Y) = m^2\left[e^{\sigma^2}-1\right] \tag{4.26}$$

4.5 A distribuição t-Student

Sejam $Z \sim N(0;1)$ e $Y \sim \chi^2(n)$ variáveis aleat'orias independentes. Então, a densidade da variável aleatória

$$U = \frac{Z}{\sqrt{Y/n}}$$

é dada por

$$f_U(u) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{u^2}{n}\right)^{-\frac{n+1}{2}} \tag{4.27}$$

Essa é a densidade t de Student, obtida por William Gosset (1876-1937), que trabalhava na Cervejaria Guinness na Irlanda. Como a cervejaria não permitia a publicação de resultados de pesquisa obtidos por seus funcionários, Gosset publicou, sob o pseudônimo de Student, o artigo "The Probable Error of a Mean" na revista Biometrika (vol. 6, no. 1).

Embora essa variável seja obtida a partir de uma função de duas outras, o resultado é uma variável aleatória unidimensional contínua e vamos estudar suas propriedades a partir da expressão 4.27. Essa expressão, certamente, é assustadora! Mas eis uma boa notícia: não precisaremos dela para calcular probabilidades! No entanto, é interessante notar duas características básicas dessa expressão: o argumento u da função aparece elevado ao quadrado e f_U depende apenas do número de graus de liberdade da qui-quadrado e, portanto, o parâmetro desta distribuição é, também, o número de graus de liberdade.

Da primeira observação resulta o fato de que f_U é simétrica em torno de zero, ou seja $f_U(u) = f_U(-u)$.

Da segunda observação resulta que cada número de graus de liberdade dá origem a uma distribuição t diferente. No entanto, pela simetria da curva, todas as distribuições t têm média 0. Além disso, o gráfico da função de densidade da t também tem forma de sino, como a distribuição normal.

Nas Figuras 4.51 a 4.54 comparam-se as distribuições t (em vermelho) com $\nu = 1, 2, 10, 30$ com a distribuição normal padrão. Nos dois gráficos superiores ($\nu = 1, 2$) fica mais nítido o fato de a distribuição t ter maior dispersão (consequência do fato de substituirmos σ pela sua estimativa s). Nos dois gráficos inferiores ($\nu = 10, 30$), o que chama a atenção é a quase coincidência da densidade t com a densidade $N(0;1)$.

Esse é um resultado importante: À medida que o número de graus de liberdade aumenta, a distribuição t de Student aproxima-se da $N(0;1)$. A variância da distribuição t com ν graus de liberdade é igual a $\frac{\nu}{\nu-2}$ $(\nu > 2)$ e podemos ver que essa variância converge a 1, que é a variância da $N(0;1)$, quando $\nu \to \infty$.

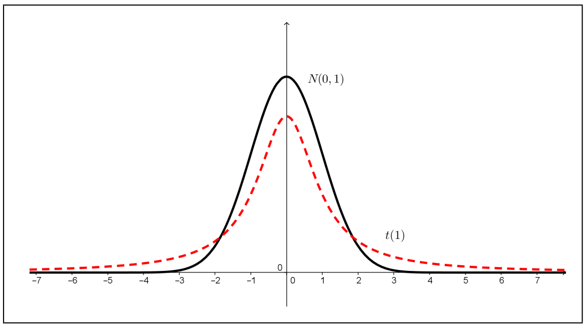


Figura 4.51 – Comparação das distribuições $N(0,1)$ e $t(1)$

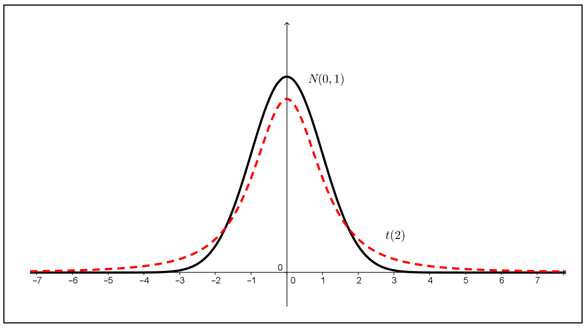


Figura 4.52 – Comparação das distribuições $N(0,1)$ e $t(2)$

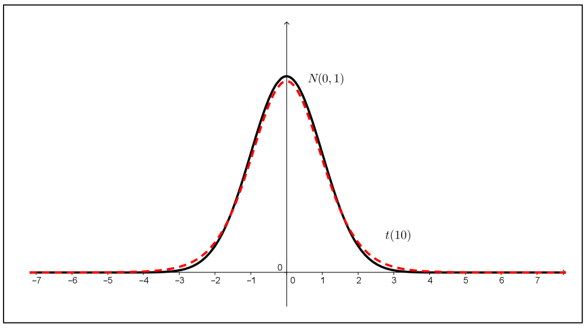


Figura 4.53 – Comparação das distribuições $N(0,1)$ e $t(10)$

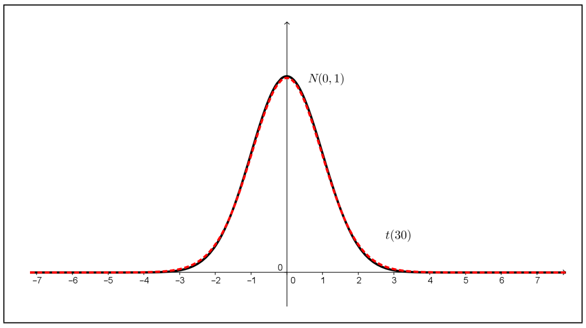


Figura 4.54 – Comparação das distribuições $N(0,1)$ e $t(30)$

A seguir, apresentam-se as propriedades da densidade $t-$Student.

DEFINIÇÃO **Densidade $t-$Student**

Sejam $Z \sim N(0;1)$ e $Y \sim \chi^2(n)$ variáveis aleatórias independentes. Se

$$T = \frac{Z}{\sqrt{Y/n}}$$

então T tem distribuição t de Student com n graus de libedade ($T \sim t_n$) e sua função de densidade é dada por

$$f_T(t) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} \qquad -\infty < t < \infty$$

Ttem-se que

$$\mathrm{E}(t_n) = 0$$

$$\mathrm{Var}(t_n) = \frac{n}{n-2} \qquad \text{se } n > 2$$

4.5.1 Tabela da t-Student

Ao contrário da distribuição normal, não existe uma relação entre as diferentes distribuições t; assim, seria necessária uma tabela para cada valor de ν. Os programas computacionais de estatística calculam probabilidades associadas a qualquer distribuição t, mas nos livros didáticos é comum apresentar uma tabela com os valores críticos de diferentes distribuições t. A razão para isso é que a maioria das aplicações da distribuição t envolve a construção de intervalos de confiança ou de testes de hipóteses.

Nessas aplicações, nosso interesse está no valor crítico associado a um nível de significância α que, como visto, é o valor da abscissa que deixa probabilidade (área) α acima dela. Vamos representar por $t_{\nu;\alpha}$ o valor crítico da distribuição $t(\nu)$. Veja a **Figura 4.55.**

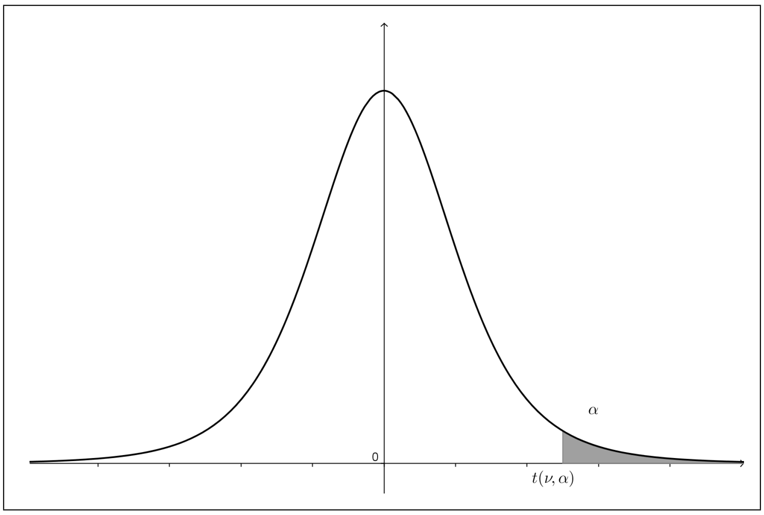


Figura 4.55 – Valor crítico $t(\nu,\alpha)$ da distribuição $t(\nu)$

A Tabela 3, apresentada no Apêndice, é uma apresentação usual dos valores críticos da distribuição t. Cada linha corresponde a um número diferente de graus de liberdade e cada coluna corresponde a uma área α na cauda superior. No corpo da tabela temos a abscissa t_α que deixa a área α acima dela, ou seja:

$$\mathrm{P}(t(\nu) > t_{\nu,\alpha}) = \alpha$$

Vamos ver, agora, exemplos de utilização da Tabela 3.

EXEMPLO 4.32 Valores críticos da t

(a) Na distribuição $t(15)$ encontre a abscissa $t_{15;0,05}$.

(b) Na distribuição $t(23)$ encontre a abscissa t tal que $\mathrm{P}(|t(23)| > t) = 0,05$.

(c) Na distribuição $t(12)$ encontre a abscissa t tal que $\mathrm{P}(|t(12)| \leq t) = 0,90$.

Solução

(a) Como o número de graus de liberdade é 15, temos de nos concentrar na linha correspondente a $gl = 15$. A abscissa $t_{0,05}$ deixa área 0,05 acima dela; logo, $t_{15;0,05} = 1,753$.

(b) Usando as propriedades da função módulo, temos a seguinte equivalência:

$$\begin{aligned}&\mathrm{P}(|t(23)| > t) = 0,05 \Longleftrightarrow\\ &\mathrm{P}(t(23) < -t) + \mathrm{P}(t(23) > t) = 0,05\end{aligned}$$

Pela simetria da densidade t, $\mathrm{P}(t(23) < -t) = \mathrm{P}(t(23) > t)$. Substituindo:

$$\begin{aligned}&\mathrm{P}(t(23) > t) + \mathrm{P}(t(23) > t) = 0,05 \Longleftrightarrow\\ &\mathrm{P}(t(23) > t) = 0,025 \Longleftrightarrow\\ &t = 2,069\end{aligned}$$

Esse último valor foi encontrado na Tabela 3, consultando-se a linha correspondente a 23 graus de liberdade e coluna correspondente à área superior de 0,025. Veja a Figura 4.56.

(c) Das propriedades da função módulo e da simetria da densidade t resultam as seguintes equivalências. Veja a Figura 4.57.

$$\begin{aligned}&\mathrm{P}(|t(12)| \le t) = 0,90 \Longleftrightarrow\\ &\mathrm{P}(-t \le t(12) \le t) = 0,90 \Longleftrightarrow\\ &\mathrm{P}(-t \le t(12) < 0) + \mathrm{P}(0 \le t(12) \le t) = 0,90 \Longleftrightarrow\\ &2 \times \mathrm{P}(0 \le t(12) \le t) = 0,90 \Longleftrightarrow\\ &\mathrm{P}(0 \le t(12) \le t) = 0,45 \Longleftrightarrow\\ &\mathrm{P}(t(12) > t) = 0,05 \Longleftrightarrow\\ &t = 1,782\end{aligned}$$

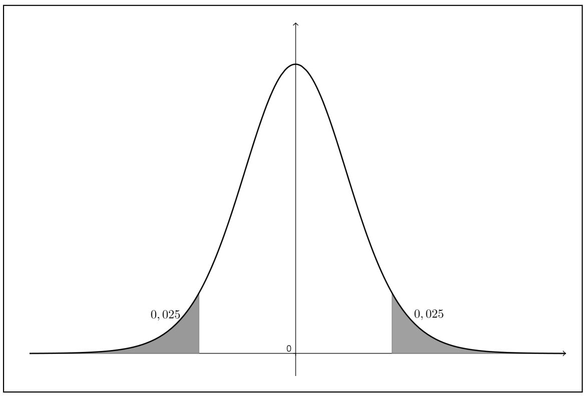


Figura 4.56 – $\mathrm{P}(|t(23)| > t) = 0,05$

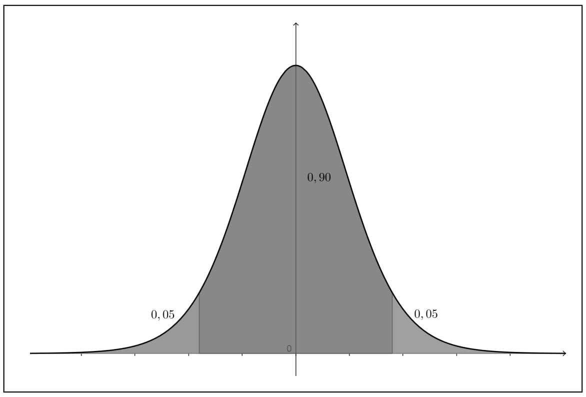


Figura 4.57 – $\mathrm{P}(|t(15)| \le t) = 0,90$

♦♦

4.6 Exercícios propostos

1. Na distribuição normal $X \sim N(\mu, \sigma^2)$, encontre:

   (a) $\Pr(X \le \mu + 2\sigma)$ (Resp.: $0,97725$)
   (b) $\Pr(|X - \mu| \le \sigma)$ (Resp.:$0,68268$)
   (c) $\Pr(|X - \mu| \le 1,96\sigma)$ (Resp.: $0,95$)
   (d) o número k tal que $\Pr(\mu - k\sigma \le X \le \mu + k\sigma) = 0,99$ (Resp.: $2,58$)
   (e) o número k tal que $\Pr(X > k) = 0,90$. (Resp.: $\mu - 1,28\sigma$)

2. Suponha que os tempos de vida de 2 marcas de aparelhos elétricos sejam variáveis aleatórias D_1 e D_2, onde $D_1 \sim N(42, 36)$ e $D_2 \sim N(45, 9)$. Se o aparelho deve ser usado por um período de 45 horas, qual marca deve ser preferida? E se for por um período de 49 horas? (Resp.: 2;1)

3. Numa distribuição normal, 31% dos elementos são menores que 45 e 8% são maiores que 64. Calcular os parâmetros que definem a distribuição. (Resp.: $\mu = 50; \sigma \approx$)

4. As vendas de um determinado produto têm distribuição aproximadamente normal com média de 500 unidades e desvio padrão de 50 unidades. Se a empresa decide fabricar 600 unidades no mês em estudo, qual a probabilidade de que não possa atender a todos os pedidos desse mês, por estar com a produção esgotada? (Resp.: $0,0228$)

5. Um produto alimentício é ensacado automaticamente, sendo o peso médio de 50 kg por saco, com desvio padrão de 1,6 kg. Os clientes exigem que, para cada saco fornecido com menos de 48 kg, o fornecedor pague uma indenização de 5 u.m..

   (a) Para 200 sacos fornecidos, qual o custo médio com indenização? (Resp.: $105,6$ u.m)

   (b) Para que o custo calculado no item anterior caia para 50 u.m., qual deveria ser a nova regulagem média da máquina? (Resp.: $50,624$)

   (c) Como o fornecedor acha que, no custo global, é desvantajoso aumentar a regulagem da máquina, ele quer comprar uma nova máquina. Qual deveria ser o desvio padrão dessa máquina para que, trabalhando com peso médio de 50 kg, em apenas 3% dos sacos se pague indenização? (Resp.: $1,064$)

6. Um teste de aptidão para o exercício de uma certa profissão exige uma sequência de operações a serem executadas rapidamente uma após a outra. Para passar no teste, o candidato deve completá-lo em, no máximo, 80 minutos. Admita que o tempo, em minutos, para completar a prova seja uma variável aleatória normal com média 90 minutos e desvio padrão 20 minutos.

   (a) Que porcentagem dos candidatos tem chance de ser aprovada? (Resp.: $30,85$)

   (b) Os 5% melhores receberão um certificado especial. Qual o tempo máximo para fazer jus a tal certificado? (Resp.: $57,2$ min)

7. O diâmetro X de rolamentos de esfera fabricados por certa fábrica tem distribuição normal com média 0,6140 e desvio padrão 0,0025. O lucro T de cada esfera depende do seu diâmetro e

   - $T = 0,10$ se a esfera é boa, isto é, $0,6100 < X < 0,6180$
   - $T = 0,05$ se a esfera é recuperável, isto é, $0,6080 < X < 0,6100$ ou $0,6180 < X < 0,6200$
   - $T = -0,10$ se a esfera é defeituosa, isto é, $X < 0,6080$ ou $X > 0,6200$

   Calcule as probabilidades de as esferas serem boas, recuperáveis e defeituosas e o lucro médio. (Resp.: $0,8904; 0,0932; 0,0164; 0,09206$)

8. Uma empresa produz televisores e garante a restituição da quantia paga se qualquer televisor apresentar algum defeito grave no prazo de 6 meses. Ela produz televisores do tipo A comum e do tipo B de luxo, com um lucro respectivo de 1000 u.m. e 2000 u.m. caso não haja restituição, e com prejuízo de 3000 u.m. e 8000 u.m., se houver restituição. Suponha que o tempo para ocorrência de algum defeito grave seja, em ambos os casos, uma variável aleatória com distribuição normal com médias 9 meses e

12 meses e desvios padrões 2 meses e 3 meses. Se tivesse que planejar uma estratégia de marketing para a empresa, você incentivaria as vendas dos aparelhos tipo A ou tipo B? (Resp.: $E(L_A) = 732,8; E(L_B) = 1772$)

9. A distribuição dos pesos de coelhos criados em uma granja pode ser representada por uma distribuição normal com média de 5 kg e desvio padrão de 0,8 kg. Um abatedouro comprará 5000 coelhos e pretende classificá-los de acordo com o peso da seguinte forma: 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 classificação? (Resp.: $4,328; 5,536; 6,024$)

10. Considere uma variável aleatória $X \sim N(3, 25)$:

    (a) Calcule $\Pr(-3 \leq X \leq 3)$ (Resp.: $0,38493$)

    (b) Calcule $\Pr(-2 \leq X \leq 8)$ (Resp.: $0,68268$)

    (c) Encontre o valor de k tal que $\Pr(X > k) = 0,05$. (Resp.: $k = 11,2$)

    (d) Encontre o valor de k tal que $\Pr(X > k) = 0,80$. (Resp.: $k = -1,2$)

11. Seja $X \sim N\left(\mu, \sigma^2\right)$. Encontre a mediana e o intervalo interquartil de X. (Resp.: $Q_2 = \mu; IQ = 1,34\sigma$)

12. O 90º percentil de uma variável aleatória $N\left(\mu, \sigma^2\right)$ é 50, enquanto o 15º percentil é 25. Encontre os valores dos parâmetros da distribuição. (Resp.: $\mu = 36,35$; $\sigma = 10,92$)

13. Uma enchedora automática enche garrafas de acordo com uma distribuição normal de média 1000 ml. Deseja-se que no máximo 1 garrafa em 100 saia com menos de 990ml. Qual deve ser o maior desvio padrão tolerável? (Resp.: $4,2918$)

14. Utilize a Tabela 4 e as propriedades da função de densidade $t-$ Student para encontrar a abscissa t que satisfaz as condições pedidas:

    (a) $P(t(18) > t) = 0,10$

    (b) $P(t(8) < t) = 0,90$

    (c) $P(t(27) < t) = 0,005$

    (d) $P(|t(30)| > t) = 0,02$

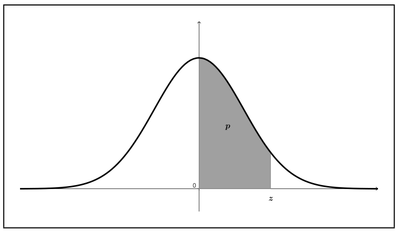
    (e) $P(|t(24)| \leq t) = 0,80$

Apêndice A

Tabelas

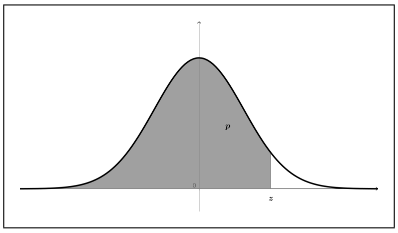
- Tabela da distribuição normal padrão – $p = \mathrm{P}(0 \leq Z \leq z)$
- Tabela da distribuição acumulada da normal padrão – $\Phi(z) = \mathrm{P}(Z \leq z),\ z \geq 0$
- Valores críticos $\chi^2_{n,\alpha}$ da qui-quadrado – $\mathrm{P}(\chi^2_n > \chi^2_{n,\alpha}) = \alpha$
- Valores críticos $t_{n;\alpha}$ da distribuição t – $\mathrm{P}(T_n > t_{n,\alpha}) = \alpha$

Tabela da distribuição normal padrão
$p = P(0 \leq Z \leq z)$



Casa inteira e 1a.decimal	2a. casa decimal									
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000
4,0	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

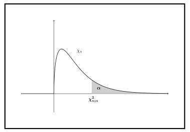
Tabela da distribuição acumulada da normal padrão
$p = \mathrm{P}(Z \leq z)$



Casa inteira	2a. casa decimal									
e 1a.decimal	0	1	2	3	4	5	6	7	8	9
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,6	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,7	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,9	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
4,0	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

Valores críticos $\chi^2_{n,\alpha}$ da qui-quadrado

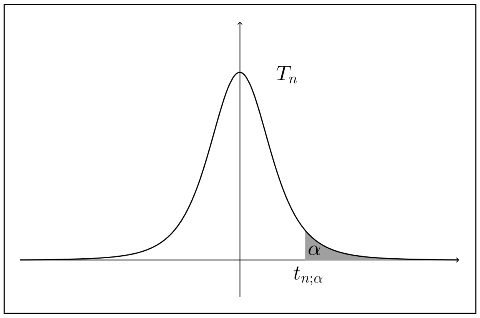
$\alpha = P(\chi^2_n > \chi^2_{n,\alpha})$



gl n	$\alpha =$ 0,999	0,995	0,990	0,980	0,975	0,950	0,900	0,800	0,200	0,100	0,050	0,025	0,020	0,010	0,005	0,001
1	0,000	0,000	0,000	0,001	0,001	0,004	0,016	0,064	1,642	2,706	3,841	5,024	5,412	6,635	7,879	10,828
2	0,002	0,010	0,020	0,040	0,051	0,103	0,211	0,446	3,219	4,605	5,991	7,378	7,824	9,210	10,597	13,816
3	0,024	0,072	0,115	0,185	0,216	0,352	0,584	1,005	4,642	6,251	7,815	9,348	9,837	11,345	12,838	16,266
4	0,091	0,207	0,297	0,429	0,484	0,711	1,064	1,649	5,989	7,779	9,488	11,143	11,668	13,277	14,860	18,467
5	0,210	0,412	0,554	0,752	0,831	1,145	1,610	2,343	7,289	9,236	11,070	12,833	13,388	15,086	16,750	20,515
6	0,381	0,676	0,872	1,134	1,237	1,635	2,204	3,070	8,558	10,645	12,592	14,449	15,033	16,812	18,548	22,458
7	0,598	0,989	1,239	1,564	1,690	2,167	2,833	3,822	9,803	12,017	14,067	16,013	16,622	18,475	20,278	24,322
8	0,857	1,344	1,646	2,032	2,180	2,733	3,490	4,594	11,030	13,362	15,507	17,535	18,168	20,090	21,955	26,124
9	1,152	1,735	2,088	2,532	2,700	3,325	4,168	5,380	12,242	14,684	16,919	19,023	19,679	21,666	23,589	27,877
10	1,479	2,156	2,558	3,059	3,247	3,940	4,865	6,179	13,442	15,987	18,307	20,483	21,161	23,209	25,188	29,588
11	1,834	2,603	3,053	3,609	3,816	4,575	5,578	6,989	14,631	17,275	19,675	21,920	22,618	24,725	26,757	31,264
12	2,214	3,074	3,571	4,178	4,404	5,226	6,304	7,807	15,812	18,549	21,026	23,337	24,054	26,217	28,300	32,909
13	2,617	3,565	4,107	4,765	5,009	5,892	7,042	8,634	16,985	19,812	22,362	24,736	25,472	27,688	29,819	34,528
14	3,041	4,075	4,660	5,368	5,629	6,571	7,790	9,467	18,151	21,064	23,685	26,119	26,873	29,141	31,319	36,123
15	3,483	4,601	5,229	5,985	6,262	7,261	8,547	10,307	19,311	22,307	24,996	27,488	28,259	30,578	32,801	37,697
16	3,942	5,142	5,812	6,614	6,908	7,962	9,312	11,152	20,465	23,542	26,296	28,845	29,633	32,000	34,267	39,252
17	4,416	5,697	6,408	7,255	7,564	8,672	10,085	12,002	21,615	24,769	27,587	30,191	30,995	33,409	35,718	40,790
18	4,905	6,265	7,015	7,906	8,231	9,390	10,865	12,857	22,760	25,989	28,869	31,526	32,346	34,805	37,156	42,312
19	5,407	6,844	7,633	8,567	8,907	10,117	11,651	13,716	23,900	27,204	30,144	32,852	33,687	36,191	38,582	43,820
20	5,921	7,434	8,260	9,237	9,591	10,851	12,443	14,578	25,038	28,412	31,410	34,170	35,020	37,566	39,997	45,315
21	6,447	8,034	8,897	9,915	10,283	11,591	13,240	15,445	26,171	29,615	32,671	35,479	36,343	38,932	41,401	46,797
22	6,983	8,643	9,542	10,600	10,982	12,338	14,041	16,314	27,301	30,813	33,924	36,781	37,659	40,289	42,796	48,268
23	7,529	9,260	10,196	11,293	11,689	13,091	14,848	17,187	28,429	32,007	35,172	38,076	38,968	41,638	44,181	49,728
24	8,085	9,886	10,856	11,992	12,401	13,848	15,659	18,062	29,553	33,196	36,415	39,364	40,270	42,980	45,559	51,179
25	8,649	10,520	11,524	12,697	13,120	14,611	16,473	18,940	30,675	34,382	37,652	40,646	41,566	44,314	46,928	52,620
26	9,222	11,160	12,198	13,409	13,844	15,379	17,292	19,820	31,795	35,563	38,885	41,923	42,856	45,642	48,290	54,052
27	9,803	11,808	12,879	14,125	14,573	16,151	18,114	20,703	32,912	36,741	40,113	43,195	44,140	46,963	49,645	55,476
28	10,391	12,461	13,565	14,847	15,308	16,928	18,939	21,588	34,027	37,916	41,337	44,461	45,419	48,278	50,993	56,892
29	10,986	13,121	14,256	15,574	16,047	17,708	19,768	22,475	35,139	39,087	42,557	45,722	46,693	49,588	52,336	58,301
30	11,588	13,787	14,953	16,306	16,791	18,493	20,599	23,364	36,250	40,256	43,773	46,979	47,962	50,892	53,672	59,703

Valores críticos $t_{n;\alpha}$ da t-Student

$\alpha = \mathrm{P}(T_n > t_{n;\alpha})$



gl	Probabilidade α na cauda superior												
n	0,150	0,100	0,060	0,050	0,040	0,030	0,025	0,020	0,010	0,005	0,0025	0,002	0,001
1	1,963	3,078	5,242	6,314	7,916	10,579	12,706	15,895	31,821	63,657	127,321	159,153	318,309
2	1,386	1,886	2,620	2,920	3,320	3,896	4,303	4,849	6,965	9,925	14,089	15,764	22,327
3	1,250	1,638	2,156	2,353	2,605	2,951	3,182	3,482	4,541	5,841	7,453	8,053	10,215
4	1,190	1,533	1,971	2,132	2,333	2,601	2,776	2,999	3,747	4,604	5,598	5,951	7,173
5	1,156	1,476	1,873	2,015	2,191	2,422	2,571	2,757	3,365	4,032	4,773	5,030	5,893
6	1,134	1,440	1,812	1,943	2,104	2,313	2,447	2,612	3,143	3,707	4,317	4,524	5,208
7	1,119	1,415	1,770	1,895	2,046	2,241	2,365	2,517	2,998	3,499	4,029	4,207	4,785
8	1,108	1,397	1,740	1,860	2,004	2,189	2,306	2,449	2,896	3,355	3,833	3,991	4,501
9	1,100	1,383	1,718	1,833	1,973	2,150	2,262	2,398	2,821	3,250	3,690	3,835	4,297
10	1,093	1,372	1,700	1,812	1,948	2,120	2,228	2,359	2,764	3,169	3,581	3,716	4,144
11	1,088	1,363	1,686	1,796	1,928	2,096	2,201	2,328	2,718	3,106	3,497	3,624	4,025
12	1,083	1,356	1,674	1,782	1,912	2,076	2,179	2,303	2,681	3,055	3,428	3,550	3,930
13	1,079	1,350	1,664	1,771	1,899	2,060	2,160	2,282	2,650	3,012	3,372	3,489	3,852
14	1,076	1,345	1,656	1,761	1,887	2,046	2,145	2,264	2,624	2,977	3,326	3,438	3,787
15	1,074	1,341	1,649	1,753	1,878	2,034	2,131	2,249	2,602	2,947	3,286	3,395	3,733
16	1,071	1,337	1,642	1,746	1,869	2,024	2,120	2,235	2,583	2,921	3,252	3,358	3,686
17	1,069	1,333	1,637	1,740	1,862	2,015	2,110	2,224	2,567	2,898	3,222	3,326	3,646
18	1,067	1,330	1,632	1,734	1,855	2,007	2,101	2,214	2,552	2,878	3,197	3,298	3,610
19	1,066	1,328	1,628	1,729	1,850	2,000	2,093	2,205	2,539	2,861	3,174	3,273	3,579
20	1,064	1,325	1,624	1,725	1,844	1,994	2,086	2,197	2,528	2,845	3,153	3,251	3,552
21	1,063	1,323	1,621	1,721	1,840	1,988	2,080	2,189	2,518	2,831	3,135	3,231	3,527
22	1,061	1,321	1,618	1,717	1,835	1,983	2,074	2,183	2,508	2,819	3,119	3,214	3,505
23	1,060	1,319	1,615	1,714	1,832	1,978	2,069	2,177	2,500	2,807	3,104	3,198	3,485
24	1,059	1,318	1,612	1,711	1,828	1,974	2,064	2,172	2,492	2,797	3,091	3,183	3,467
25	1,058	1,316	1,610	1,708	1,825	1,970	2,060	2,167	2,485	2,787	3,078	3,170	3,450
26	1,058	1,315	1,608	1,706	1,822	1,967	2,056	2,162	2,479	2,779	3,067	3,158	3,435
27	1,057	1,314	1,606	1,703	1,819	1,963	2,052	2,158	2,473	2,771	3,057	3,147	3,421
28	1,056	1,313	1,604	1,701	1,817	1,960	2,048	2,154	2,467	2,763	3,047	3,136	3,408
29	1,055	1,311	1,602	1,699	1,814	1,957	2,045	2,150	2,462	2,756	3,038	3,127	3,396
30	1,055	1,310	1,600	1,697	1,812	1,955	2,042	2,147	2,457	2,750	3,030	3,118	3,385
31	1,054	1,309	1,599	1,696	1,810	1,952	2,040	2,144	2,453	2,744	3,022	3,109	3,375
32	1,054	1,309	1,597	1,694	1,808	1,950	2,037	2,141	2,449	2,738	3,015	3,102	3,365
33	1,053	1,308	1,596	1,692	1,806	1,948	2,035	2,138	2,445	2,733	3,008	3,094	3,356
34	1,052	1,307	1,595	1,691	1,805	1,946	2,032	2,136	2,441	2,728	3,002	3,088	3,348
35	1,052	1,306	1,594	1,690	1,803	1,944	2,030	2,133	2,438	2,724	2,996	3,081	3,340
36	1,052	1,306	1,593	1,688	1,802	1,942	2,028	2,131	2,434	2,719	2,990	3,075	3,333
37	1,051	1,305	1,592	1,687	1,800	1,940	2,026	2,129	2,431	2,715	2,985	3,070	3,326
38	1,051	1,304	1,591	1,686	1,799	1,939	2,024	2,127	2,429	2,712	2,980	3,064	3,319
39	1,050	1,304	1,590	1,685	1,798	1,937	2,023	2,125	2,426	2,708	2,976	3,059	3,313
40	1,050	1,303	1,589	1,684	1,796	1,936	2,021	2,123	2,423	2,704	2,971	3,055	3,307