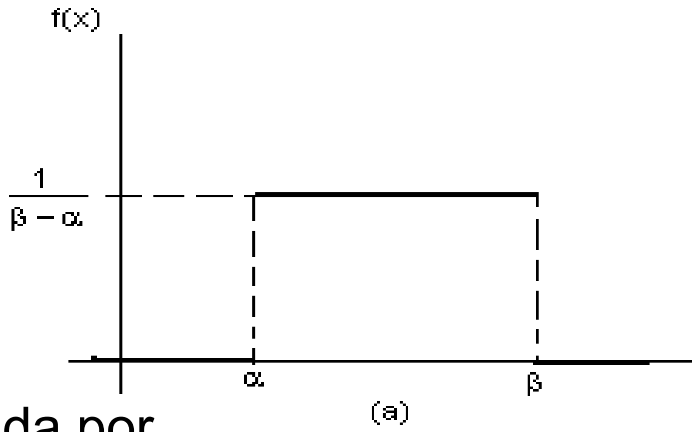
5. PRINCIPAIS MODELOS CONTÍNUOS

2010

5.1. Modelo uniforme

Uma v.a. contínua X tem distribuição uniforme com parâmetros α e β ($\alpha < \beta$) se sua função densidade de probabilidade é dada por

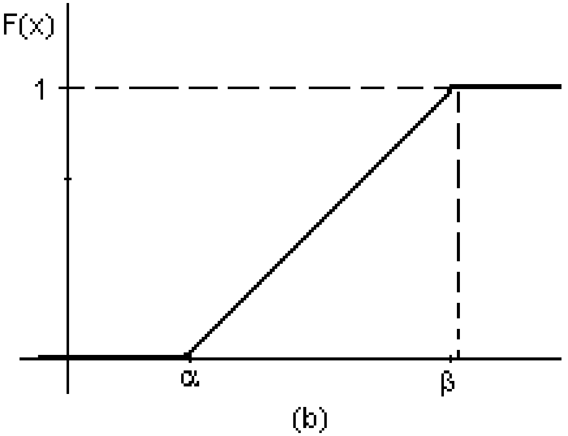
$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha}, & \alpha \le x \le \beta \\ 0, & \text{c.c.} \end{cases}$$



Notação: $X \sim U(\alpha, \beta)$.

A função de distribuição acumulada é dada por

$$F(x) = \begin{cases} 0, & \text{se} \quad x < \alpha, \\ \dfrac{x-\alpha}{\beta - \alpha}, & \text{se} \quad \alpha \le x \le \beta, \\ 1, & \text{se} \quad x > \beta. \end{cases}$$



Propriedades:

$$\mathrm{E}(X) = \frac{\alpha + \beta}{2} \quad \text{e} \quad \mathrm{Var}(X) = \frac{(\alpha - \beta)^2}{12}.$$

Exemplo

A dureza de uma peça de aço pode ser pensada como sendo uma variável aleatória uniforme no intervalo (50,70) unidades. Qual a probabilidade de que uma peça tenha dureza entre 55 e 60?

Solução. X representa a dureza de uma peça de aço, sendo que X ~ U(50, 70) e

$$f(x) = \begin{cases} \dfrac{1}{20}, & 50 \leq x \leq 70, \\ 0, & c.c. \end{cases}$$

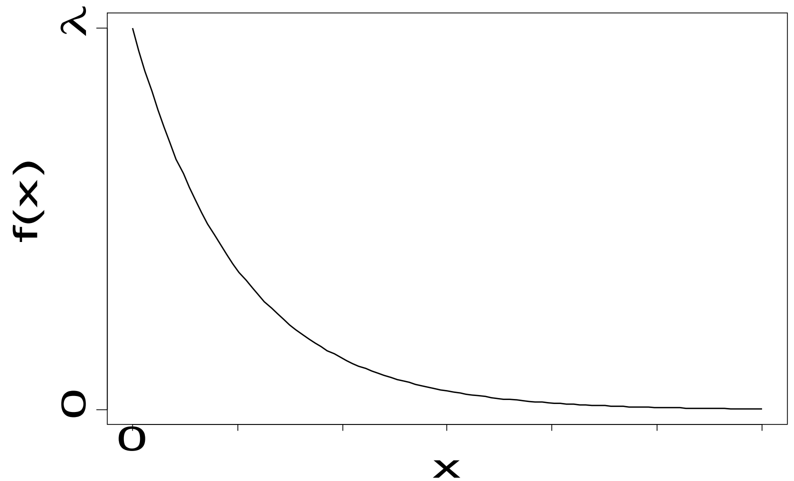
Portanto,

$$P(55 < X < 60) = \int_{55}^{60} \frac{1}{20} dx = \frac{5}{20} = 0,25.$$

5.2. Modelo exponencial

Uma v.a. contínua X tem distribuição exponencial com parâmetro $\lambda > 0$ se sua função de densidade é dada por

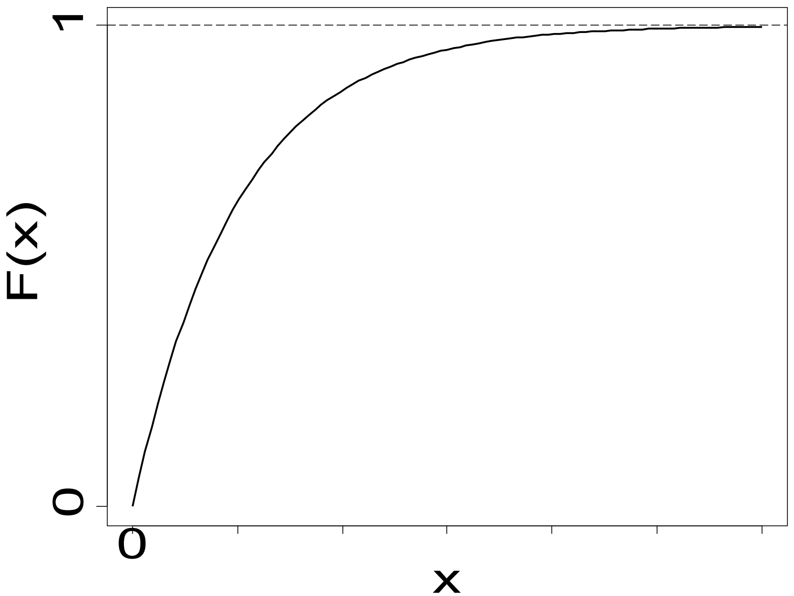
$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & \text{c.c.} \end{cases}$$



Notação: $X \sim Ex(\lambda)$.

A função de distribuição acumulada é dada por

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & \text{c.c.} \end{cases}$$



Propriedades:

$$E(X) = 1/\lambda \quad \text{e} \quad Var(X) = 1/\lambda^2.$$

5.2. Modelo exponencial

Propriedade. Se X ~ Ex(λ), então P(X > a + b| X > b) = P(X > a).

É a única distribuição contínua com esta propriedade ("falta de memória").

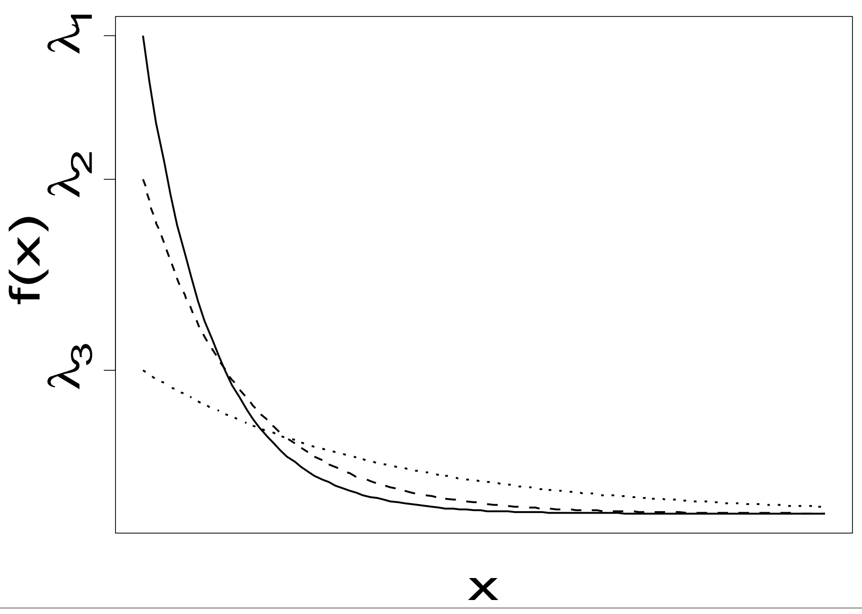
Observação. Também encontramos X ~ Ex(α), em que

$$f(x) = \begin{cases} \dfrac{1}{\alpha} e^{-\frac{x}{\alpha}}, & x \geq 0, \\ 0, & \text{c.c.} \end{cases}$$

Relação: $\alpha = 1 / \lambda$.

α: escala e λ: taxa.

Exemplo. Diferentes valores de λ.



Exemplo

O intervalo de tempo entre a passagem de veículos por um ponto em uma estrada tem distribuição exponencial com média de 0,55 minutos. Qual a probabilidade de que o tempo seja no máximo 12 segundos?

Solução. Se X é o intervalo de tempo entre a passagem de veículos, temos E(X) = 0,55 min, α = E(X) = 0,55 e X ~ Ex(α = 0,55). Ou seja,

$$F(x) = \begin{cases} 1 - e^{-\frac{x}{0,55}}, & x \geq 0, \\ 0, & \text{c.c.} \end{cases}$$

$$\mathrm{P}(X \leq 12\, s) = \mathrm{P}(X \leq 12/60 \min) = 1 - e^{-\frac{0,2}{0,55}} = 0{,}305.$$

Em Excel: =DISTEXPON(0,2; 1/0,55; VERDADEIRO)

5.3. Modelo normal (ou gaussiano)

Uma variável aleatória contínua X tem distribuição normal com média μ e variância σ^2 se sua função densidade é dada por

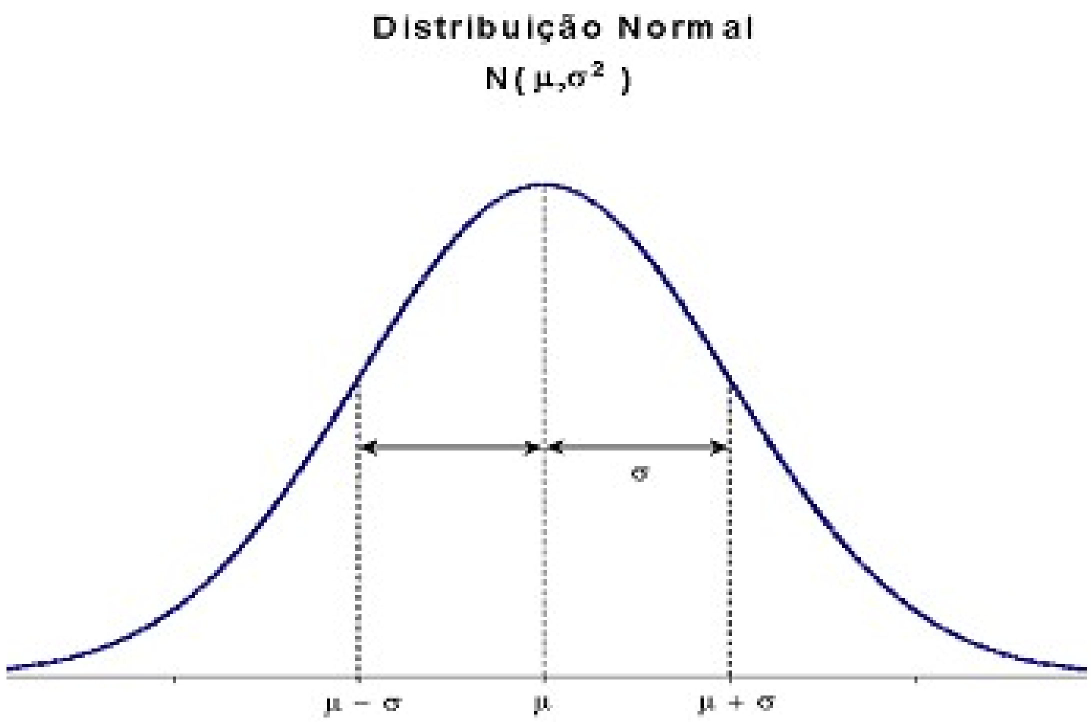
$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\left(\frac{x-\mu}{\sigma}\right)^2}, \; x \in R.$$

Notação: $X \sim N(\mu, \sigma^2)$.

Propriedades: $E(X) = \mu$,

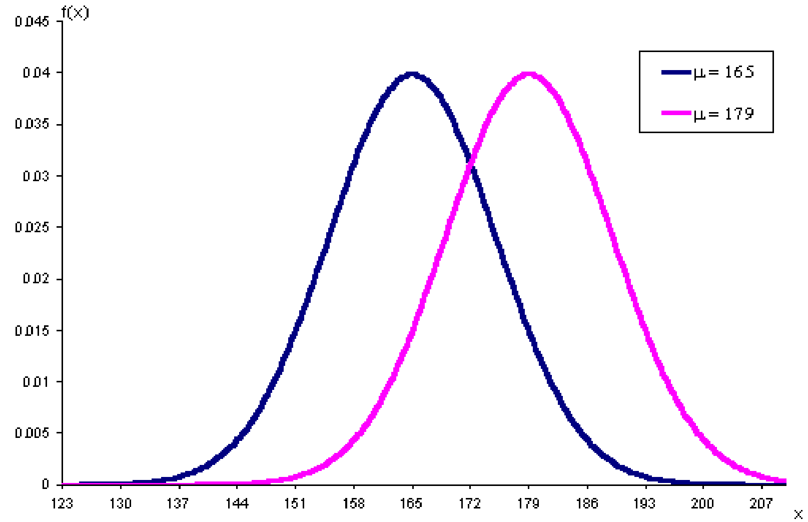
$Var(X) = \sigma^2$

e mediana = moda = μ.

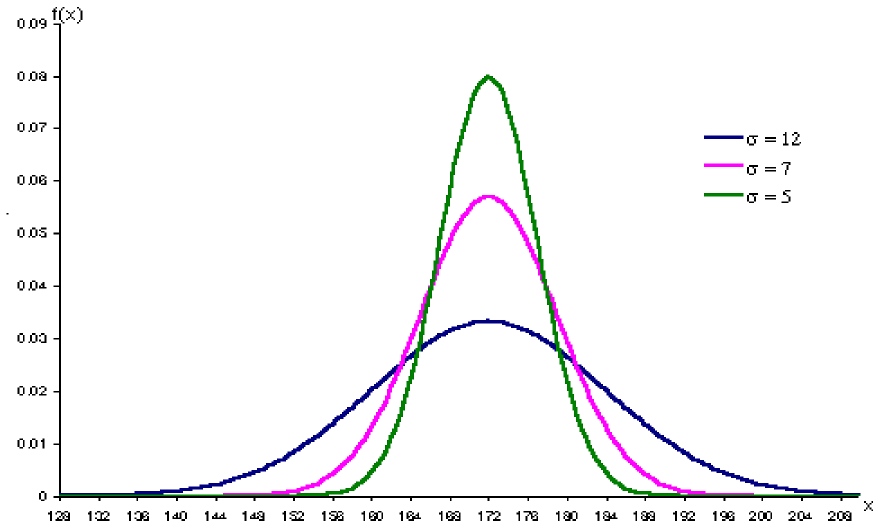


Exemplos

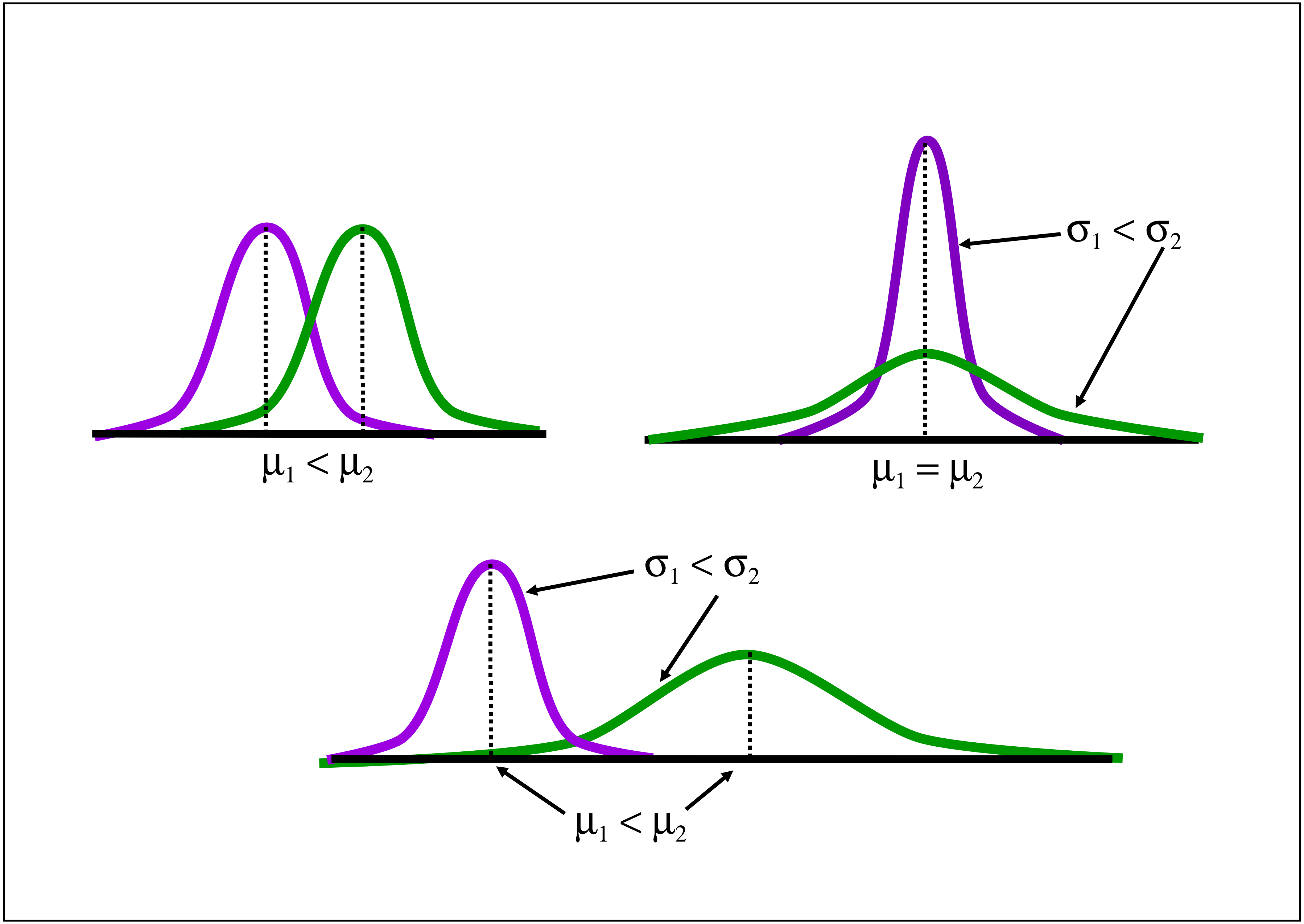
Distribuições normais com médias diferentes e variâncias iguais.



Distribuições normais com médias iguais e variâncias diferentes.



Exemplos

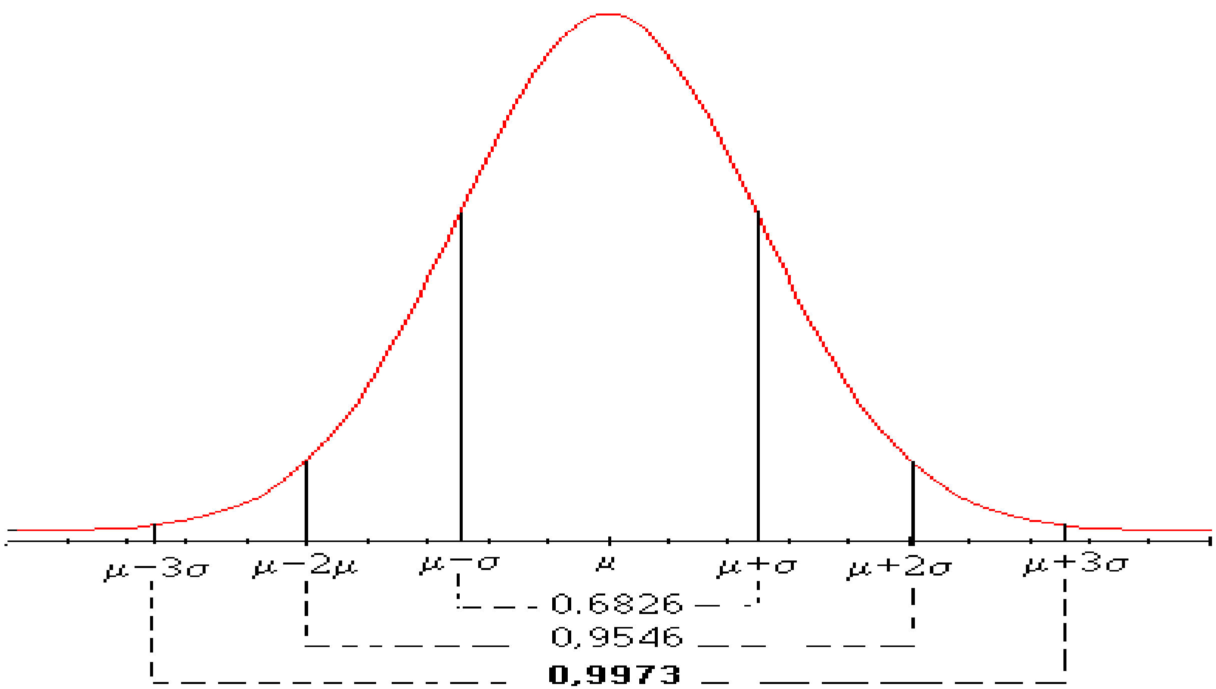


Propriedades

(a) A distribuição é simétrica em relação à média.

(b) Como a área total sob curva é igual a 1, à esquerda e à direita de μ a área é igual a 0,5.

(c)
$$P(\mu - \sigma \leq X \leq \mu + \sigma) = 0{,}6896,$$
$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = 0{,}9546 \text{ e}$$
$$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) = 0{,}9973.$$



Propriedades

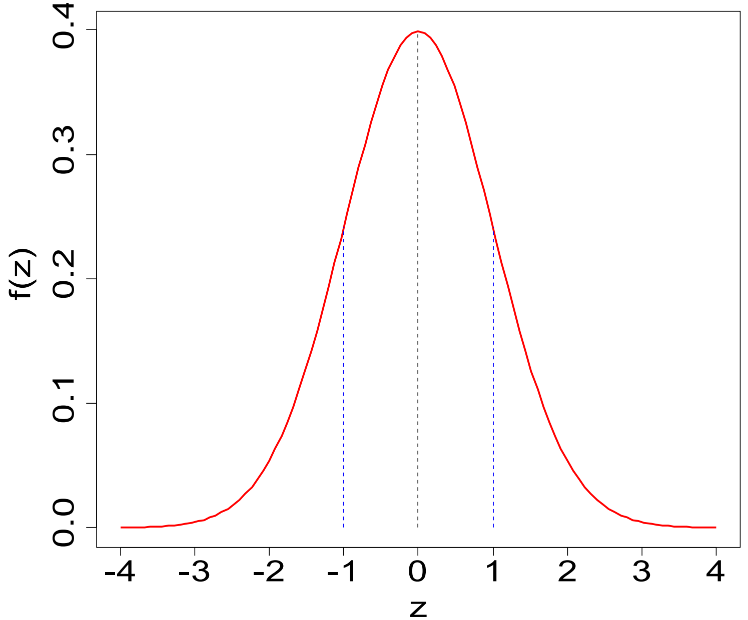
A função de distribuição acumulada de uma v.a. X ~ N(μ, σ^2) é

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{ -\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2 \right\} dt.$$

Integral sem solução analítica. Cálculo de probabilidades com o auxílio de tabelas.

Normal padrão ou reduzida. Se Z é uma v.a. normal com média 0 e variância 1, então Z é chamada de uma v.a. normal padrão ou reduzida e sua função densidade é

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \; z \in R.$$



A função de distribuição acumulada de uma v.a. Z ~ N(0,1) é

$$\Phi(z) = P(Z \le z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}t^2)dt.$$

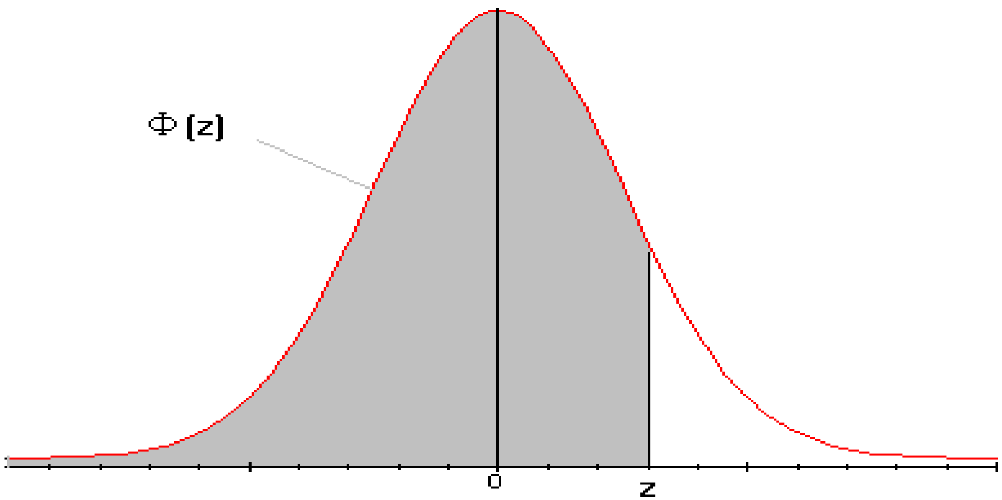
Uso da tabela normal

Table A.3. Areas under the normal curve.

Z ~ N(0,1): distribuição normal padrão.

Valores no corpo da tabela: $\Phi(z) = P(Z \leq z)$, z com duas decimais.

$$\Phi(z) = P(Z \leq z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}t^2)dt, \; -3{,}40 \leq z \leq 3{,}49.$$



Uso da tabela normal

1ª coluna: parte inteira de z e 1ª decimal.

1ª linha: 2ª decimal de z.

Exemplo. $P(Z \leq -1,25)$ é encontrada na interseção da linha correspondente a –1,2 com a coluna 0,05:

2ª decimal

Parte inteira e 1ª decimal

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-3,4										
...										
-1,2						0,1056				
...										
3,4										

Resposta. $P(Z \leq -1,25) = 0,1056$.

Exemplo

Se Z ~ N(0,1), calcule
(a) P(Z < 1,80),
(b) P(0,80 < Z < 1,40),
(c) P(Z > -0,57) e
(d) o valor de k tal que
P(Z < k) = 0,05.

Em Excel:

(a) =DIST.NORMP(1,8).

(b) = DIST.NORMP(1,4) – DIST.NORMP(0,8).

(c) =1-DIST.NORMP(-0,57).

(d) =INV.NORMP(0,05).

Solução. Da tabela normal padrão tem-se

$$\text{(a)}\ \mathrm{P}(Z < 1{,}80) = \Phi(1{,}80) = 0{,}9641,$$

$$\text{(b)}\ \mathrm{P}(0{,}80 < Z < 1{,}40) = \Phi(1{,}40) - \Phi(0{,}80) = 0{,}9192 - 0{,}7881 = 0{,}1311,$$

$$\text{(c)}\ \mathrm{P}(Z > -0{,}57) = 1 - \mathrm{P}(Z \le -0{,}57) = 1 - 0{,}2843 = 0{,}7157,$$

$$\text{(d)}\ \mathrm{P}(Z < k) = 0{,}05 \Rightarrow k = -1{,}64.$$

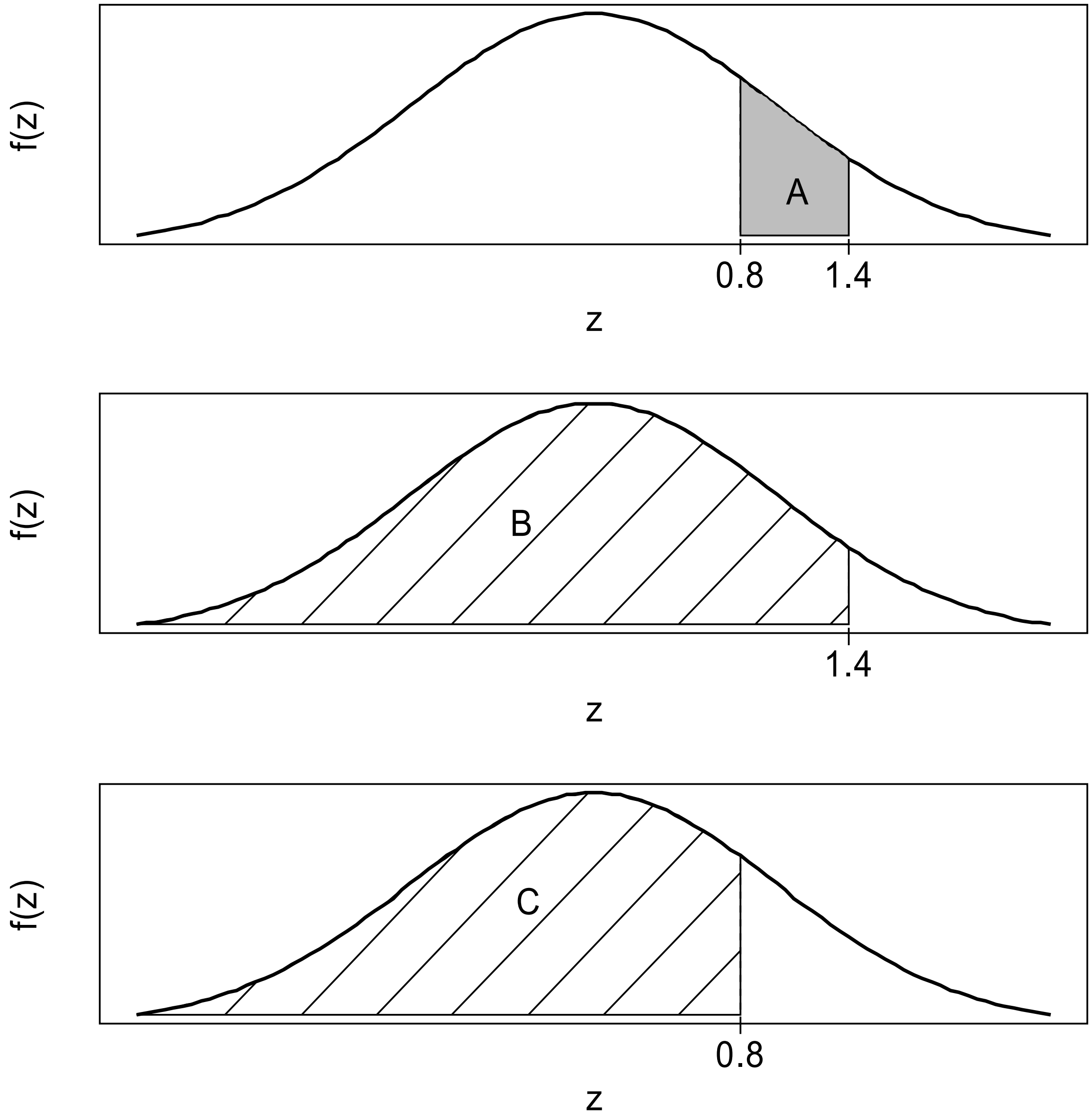
Observação. Para todo k > 0,

$$(i)\ \mathrm{P}(Z \le -k) = 1 - \mathrm{P}(Z \le k) \quad \text{e}$$

$$(ii)\ \mathrm{P}(-k \le Z \le k) = 2\mathrm{P}(Z \le k) - 1 = 1 - 2\mathrm{P}(Z \le -k).$$

Exemplo (b)

A = B – C, sendo que B e C são encontradas na tabela normal.



Transformação linear de uma variável normal

Se $X \sim N(\mu, \sigma^2)$, então $Y = a + bX \sim N(\mu_Y, \sigma_Y^2)$, sendo que $\mu_Y = a + b\mu$ e $\sigma_Y^2 = b^2 \sigma^2$.

Tomando $a = -\mu / \sigma$ e $b = 1 / \sigma$ obtemos a padronização

$$Z = \frac{X - \mu}{\sigma} \sim \mathrm{N}(0,1).$$

Distribuição normal padrão ou reduzida.

Exemplo. Se $X \sim N(90,100)$, determinar

- $P(80 < X < 100)$,
- $P(|X - 90| < 30)$ e
- o valor de a tal que $P(90 - 2a < X < 90 + 2a) = 0,99$.

Exemplo

$$(a)\,\mathrm{P}(80 < X < 100) = \mathrm{P}\left(\frac{80-90}{10} < \frac{X-\mu}{\sigma} < \frac{100-90}{10}\right) = \mathrm{P}(-1{,}00 < Z < 1{,}00)$$

$$= 2\mathrm{P}(Z \le 1{,}00) - 1 = 2 \times 0{,}8413 - 1 = 0{,}6826.$$

$$(b)\ \mathrm{P}(|X-90| < 30) = \mathrm{P}(-30 < X - 90 < 30) = \mathrm{P}\left(-\frac{30}{10} < \frac{X-90}{10} < \frac{30}{10}\right)$$

$$= \mathrm{P}(-3{,}00 < Z < 3{,}00) = 2\mathrm{P}(Z < 3{,}00) - 1$$

$$= 2 \times 0{,}9987 - 1 = 0{,}9974.$$

$$(c)\,\mathrm{P}(90 - 2a < X < 90 + 2a) = \mathrm{P}(-2a < X - 90 < 2a) = \mathrm{P}\left(-\frac{2a}{10} < \frac{X-90}{10} < \frac{2a}{10}\right)$$

$$= 2\mathrm{P}\left(Z \le \frac{a}{5}\right) - 1 = 0{,}99 \Rightarrow \mathrm{P}\left(Z < \frac{a}{5}\right) = 0{,}995$$

$$\Rightarrow \frac{a}{5} = 2{,}57 \Rightarrow a = 12{,}85.$$

Exercício. Resolver o exemplo utilizando o Excel.

Exemplo

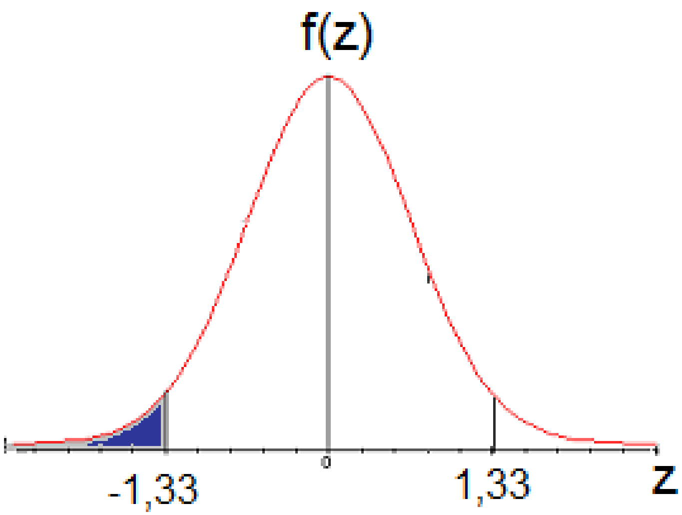
A resistência à compressão de blocos de concreto tem distribuição normal com média 120 N/mm^2 e desvio padrão 15 N/mm^2.

- Qual a probabilidade de que a resistência seja inferior a 100 N/mm^2?

Solução. Definimos X como a resistência dos blocos à compressão. Pelo enunciado, $X \sim N(120, 15^2)$. Calculamos

$$P(X < 100) = P\left(Z < \frac{100 - 120}{15}\right) = P(Z \leq -1{,}33)$$

$$= \Phi(-1{,}33) = 0{,}0918.$$



Em Excel:

= DIST.NORM(100; 120; 15;VERDADEIRO)

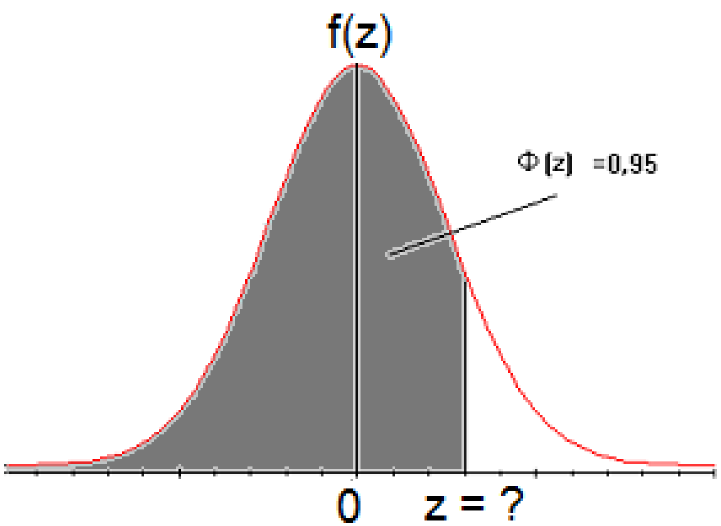
Exemplo

(b) Qual a resistência correspondente a 95% dos blocos de menor resistência?

Solução. Devemos encontrar x tal que $P(X \leq x) = 0{,}95$. Após uma transformação,

$$P(X \leq x) = P\left(Z \leq \frac{x - 120}{15} \right) = 0{,}95.$$

Iniciamos encontrando z tal que $\Phi(z) = 0{,}95$.



Da tabela normal, z = 1,64. Logo, x = 120 + 1,64 × 15 = 144,6 N/mm².

Em Excel:

= INV.NORM(0,95; 120; 15)
= 144,6728.

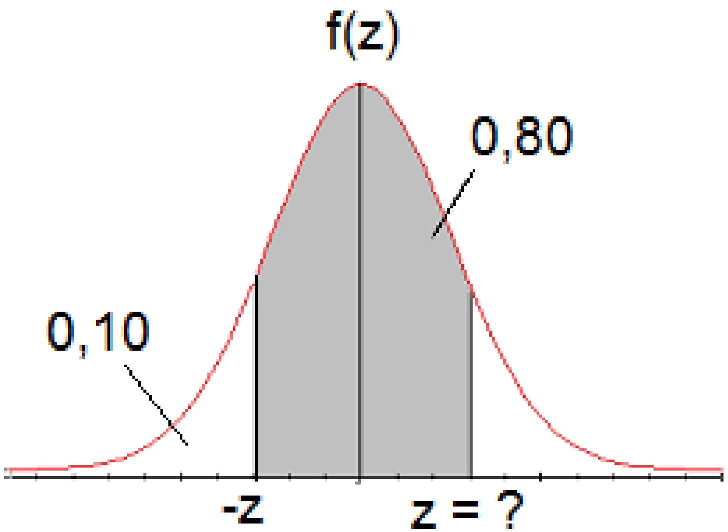
Obs. Este valor é o quantil (ou percentil) 95% da distribuição.

Exemplo

(c) Qual o intervalo central correspondente a 80% de todos os valores da resistência?

Solução. Devemos encontrar x_1 e x_2 tais que

$$P(x_1 \le X \le x_2) = 0{,}80 \Rightarrow P\left(\frac{x_1 - 120}{15} \le Z \le \frac{x_2 - 120}{15}\right) = 0{,}80.$$



Probabilidade acumulada até o ponto z é igual a 0,90.

Procuramos z tal que $\Phi(z) = 0{,}90$.
Da tabela normal, z = 1,28. Logo,

$$\frac{x_1 - 120}{15} = -1{,}28 \Rightarrow x_1 = 120 - 15 \times 1{,}28 = 100{,}8 \text{ N/mm}^2,$$

$$\frac{x_2 - 120}{15} = 1{,}28 \Rightarrow x_2 = 120 + 15 \times 1{,}28 = 139{,}2 \text{ N/mm}^2.$$

Em Excel:

x_1 = INV.NORM(0,10; 120; 15) = 100,7767

e x_2 = INV.NORM(0,90; 120; 15) = 139,22326.

A escala sigma

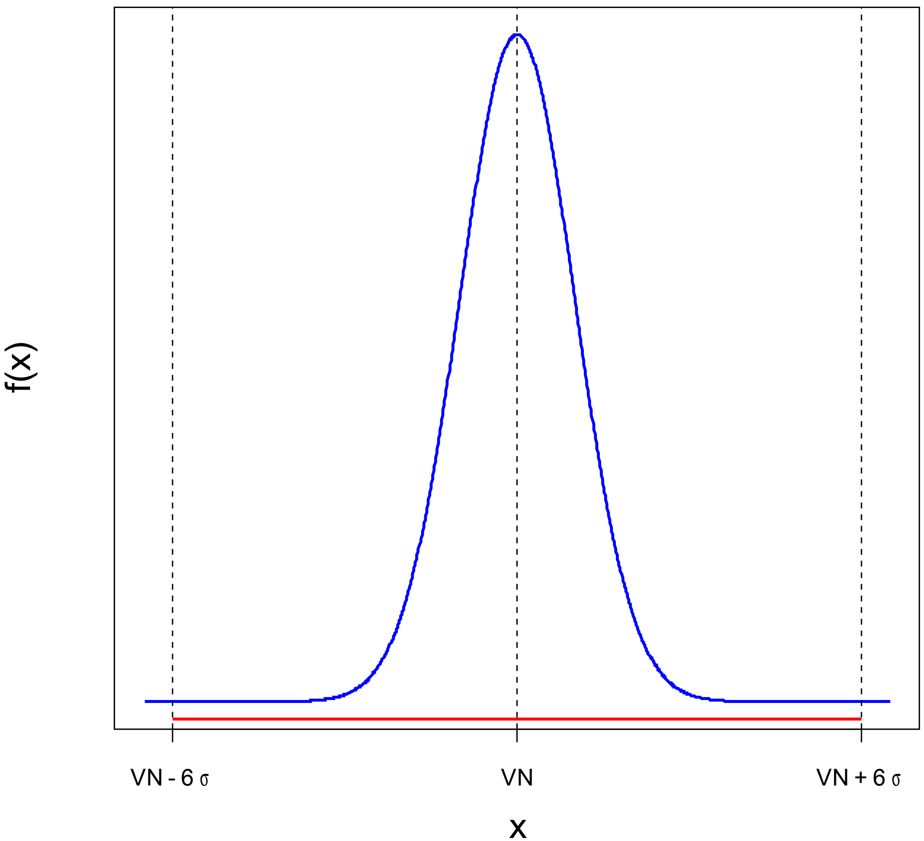
Utilizada para medir o nível de qualidade de um processo de produção.

Quanto maior o número de sigmas (σ), melhor.

X representa uma característica de um item, sendo que $X \sim N(VN, \sigma^2)$.

μ = VN = valor nominal.

Limites de especificação: LIE = VN – 6σ e LSE = VN + 6σ.



$P(X < VN - 6\sigma) + P(X > VN + 6\sigma)$
$= 2\, P(X < VN - 6\sigma)$
$= 2 \times 9{,}865876 \times 10^{-10}$
$= 1{,}973175 \times 10^{-9}$.

Em Excel: =2*DIST.NORMP(-6)
$= 1{,}98024 \times 10^{-9}$.

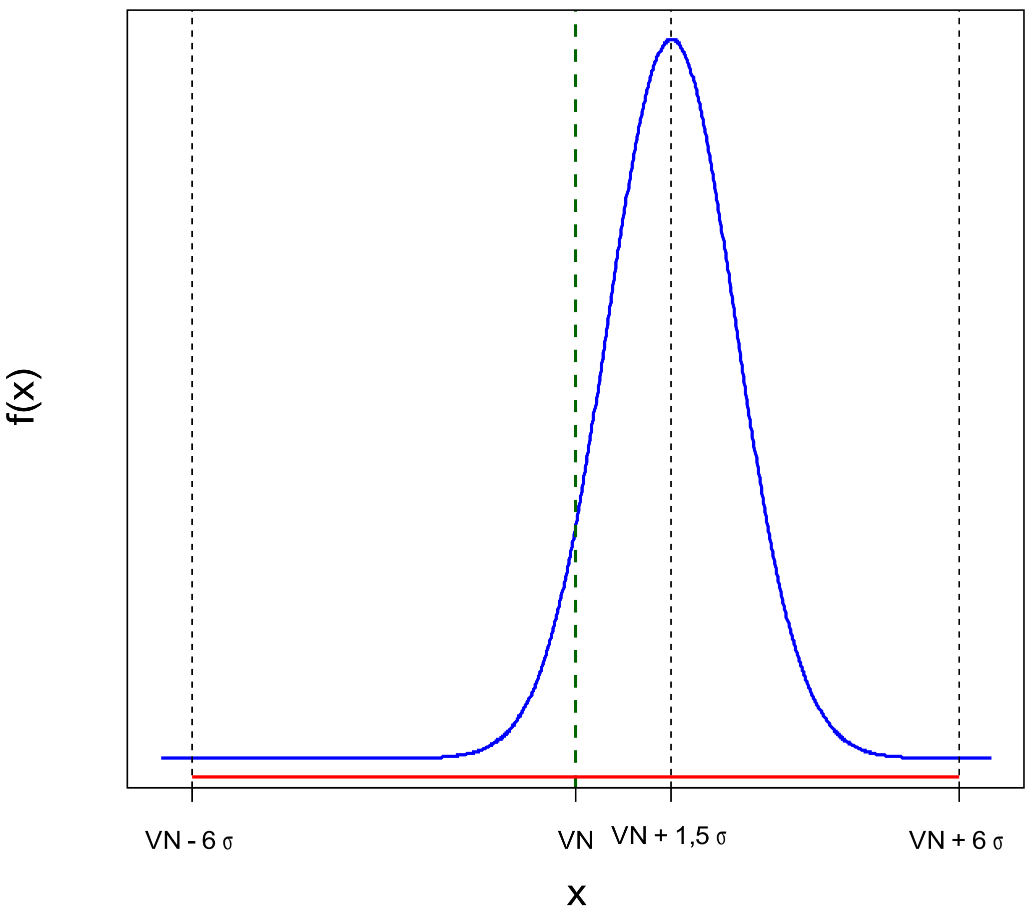
Corresponde, em média, a cerca de dois itens que não atendem às especificações a cada bilhão de itens produzidos.

=2*DIST.NORMP(-6)* 1E9 = 1,980.

A escala sigma

O processo sofre uma alteração. A média passa a ser μ = VN – 1,5σ ou μ = VN + 1,5σ.

Considere $X \sim N(\mu, \sigma^2)$, em que μ = VN + 1,5σ.



P(X < VN – 6σ) + P(X > VN + 6σ) = 3,397673 $\times$ 10^{-6}.

Em Excel:
=DIST.NORM(-6;1,5;1; VERDADEIRO) + 1– DIST.NORM(6;1,5;1; VERDADEIRO) = 3,4008 $\times$ 10^{-6}.

Corresponde, em média, a cerca de 3,4 itens que não atendem às especificações a cada milhão de itens produzidos.

A escala sigma

Nível	Média de defeitos por milhão
2σ	308537
3 σ	66807
4 σ	6210
5 σ	233
6 σ	3,4

4 σ	6 σ
Sete horas de falta de energia por mês	Uma hora de falta de energia a cada 34 anos
5000 cirurgias incorretas por semana	1,7 cirurgia incorreta por semana
15 minutos de fornecimento de	Um minuto de fornecimento de
água não potável por dia	água não potável a cada sete meses
Fonte: Keene, S. (2000), *Reliability Review* 20, p.19.	

Propriedade

Se $X_1,\ldots,X_n$ são v.a. independentes tais que $X_i \sim N(\mu, \sigma^2)$, para i = 1,...,n, então, a v.a.

$$Y = X_1 + \cdots + X_n = \sum_{i=1}^{n} X_i$$

é tal que $Y \sim N(n\mu, n\sigma^2)$.

Padronização:

$$Z = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma} = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0,1).$$

Exemplo. O peso de uma caixa de peças é uma v.a. normal com média 65 kg e desvio padrão de 4 kg. Um carregamento de 120 caixas de peças é despachado. Qual a probabilidade de que a carga pese entre 7.893 kg e 7.910 kg?

Exemplo

Solução. Pelo enunciado,

X_i : peso da i-ésima caixa $\Rightarrow$ $X_i \sim \mathrm{N}(65{,}16),\ i = 1,\cdots,120.$

Logo,

Y : peso da carga $\Rightarrow$ $Y = \sum_{i=1}^{120} X_i \sim \mathrm{N}(120\times 65{,}120\times 16),$

$Y \sim \mathrm{N}(7800{,}1920).$

Calculamos

$$\mathrm{P}(7893 \le Y \le 7910) = \mathrm{P}\left(\frac{7893-7800}{\sqrt{1920}} \le Z \le \frac{7910-7800}{\sqrt{1920}}\right)$$

$$= \mathrm{P}(2{,}12 \le Z \le 2{,}51) = \Phi\,(2{,}51) - \Phi\,(2{,}12)$$

$$= 0{,}4940 - 0{,}4830 = 0{,}0110.$$

Teorema central do limite

Se $X_1, X_2, ..., X_n$ é uma amostra aleatória de tamanho n de uma distribuição com média μ e desvio padrão σ ($0 < \sigma < \infty$), então a distribuição aproximada de

$$Z = \frac{\sqrt{n}(\overline{X} - \mu)}{\sigma}$$ é normal padrão N(0,1),

sendo que $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ é a média amostral.

Observações.

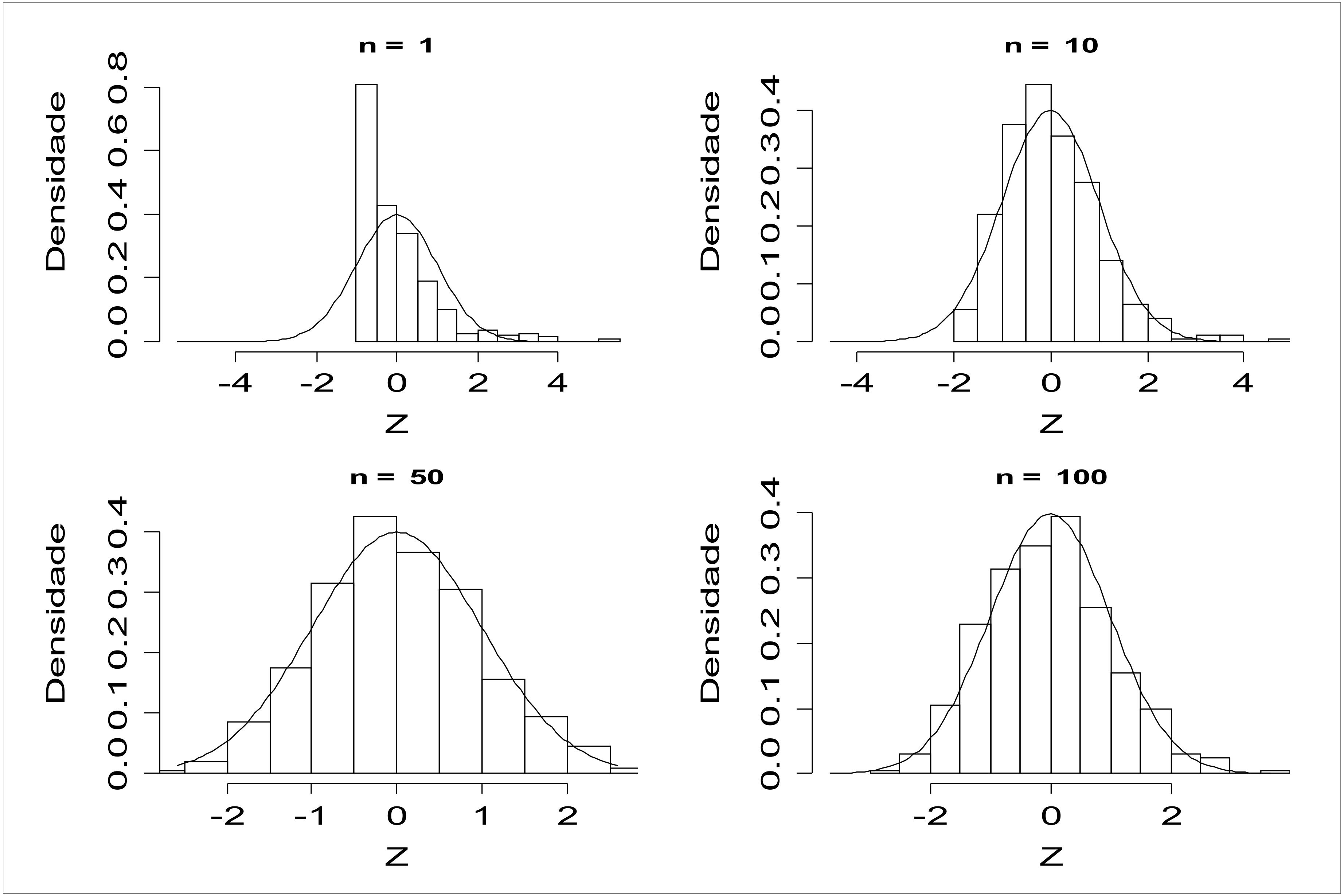
(1) Quanto maior n, melhor a aproximação.

(2) A distribuição das variáveis X pode ser discreta ou contínua.

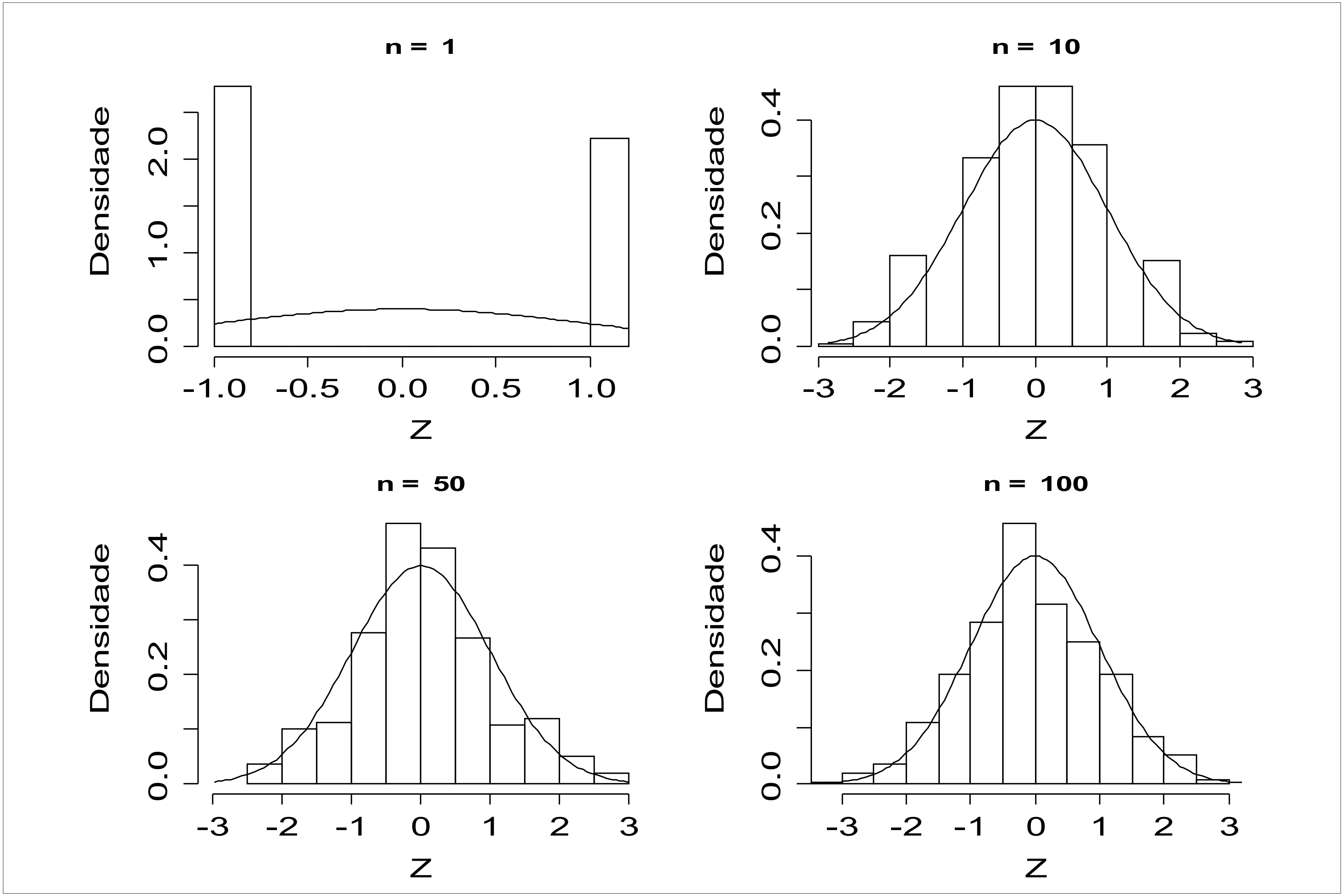
(3) A distribuição aproximada de

$$\sum_{i=1}^{n} X_i \text{ é } \mathrm{N}(n\mu, n\sigma^2).$$

Teorema central do limite – Distribuição exponencial



Teorema central do limite – Distribuição Bernoulli (p = 0,45)



Exemplo

Após arredondamento para o inteiro mais próximo, 48 números são somados. Os erros de arredondamento individuais são uniformemente distribuídos no intervalo (-0,5; 0,5). Qual a probabilidade de que a soma dos números arredondados seja diferente da verdadeira soma por mais de 3 unidades (em ambos os sentidos) ?

Solução. Utilizando o teorema central do limite obtemos uma solução aproximada.

X_i, i = 1,...,48 são os erros de arredondamento tais que $X_i \sim U(-0,5; 0,5)$,

$E(X_i) = (-0,5 + 0,5) / 2 = 0$ e $Var(X_i) = [0,5 – (-0,5)]^2 / 12 = 1 / 12$ (veja lâmina 2).

O erro de arredondamento E é dado por $E = X_1 + X_2 + \ldots + X_{48}$, sendo que a distribuição aproximada é $E \sim N(48 \times 0,\ 48 \times 1/12) = N(0,4)$.

Devemos calcular $P((E < -3) \cup (E > 3))$, que é igual a $P(E < -3) + P(E > 3)$.

Usando a distribuição aproximada, $P(E < -3) + P(E > 3) = 2\ P(E < -3)$

$$= 2P\left(\frac{E-0}{2} < \frac{-3-0}{2}\right) = 2P(Z < -1,50) = 2 \times 0,0668 = 0,1336.$$

5.4. Modelo de Weibull

Uma variável aleatória contínua X tem distribuição de Weibull com parâmetros de escala $\alpha > 0$ e forma $\beta > 0$ se sua função densidade é dada por

$$f(x) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta - 1} e^{-\left(\frac{x}{\alpha}\right)^{\beta}}, \quad x \geq 0.$$

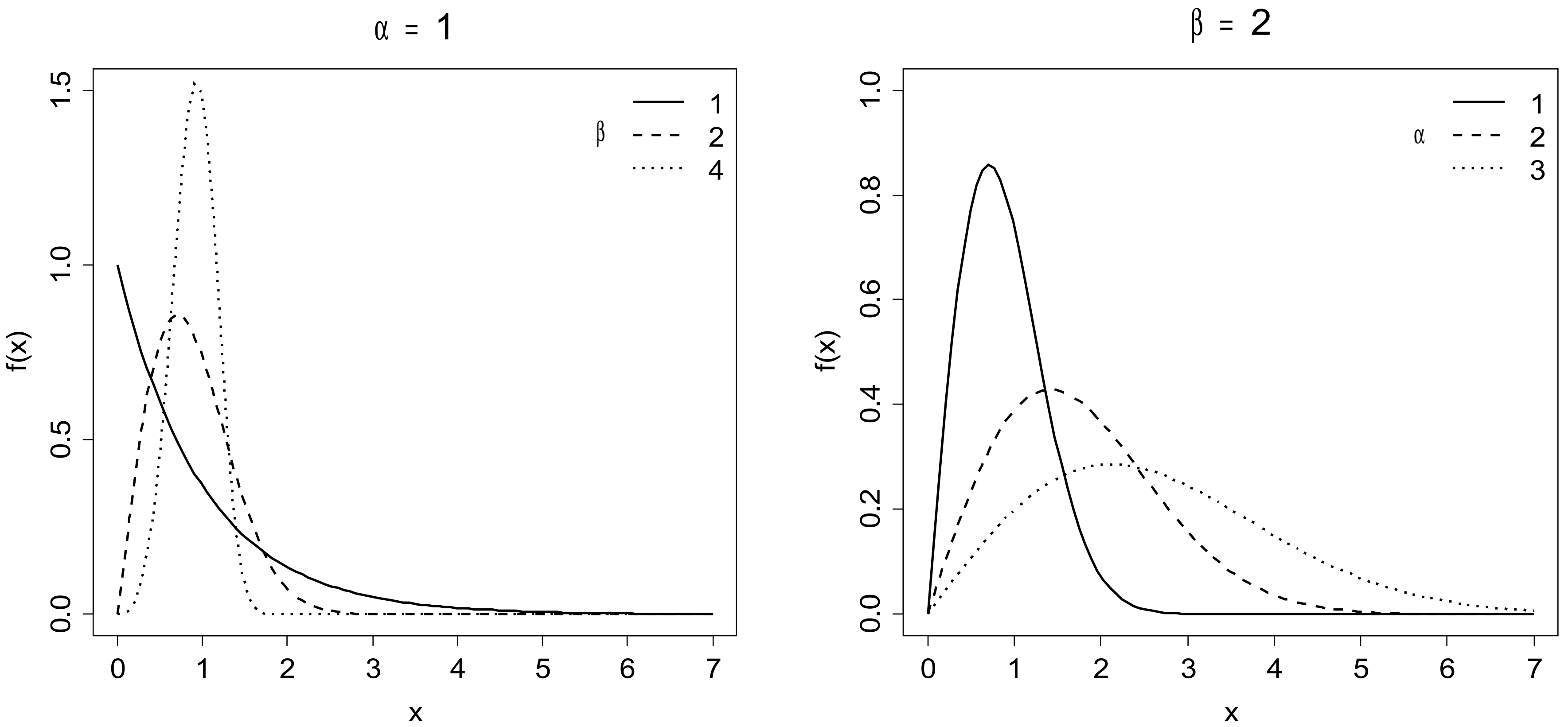
Função distribuição acumulada:

$$F(x) = P(X \leq x) = 1 - e^{-\left(\frac{x}{\alpha}\right)^{\beta}}, \quad x \geq 0.$$

Notação: $X \sim W(\alpha, \beta)$.

Obs. Se $\beta = 1$, $X \sim Ex(\alpha)$ (lâmina 5).

Exemplos



Outras distribuições contínuas: gama, Gumbel, log-normal, Fréchet, Pearson, valor extremo, etc.

Período de retorno

Eventos extremos: cheias, secas, velocidade do vento, etc.

X: vazão máxima anual em uma certa estação de medição (v. a. contínua).

X_i: vazão máxima ocorrida no ano i, i = 1, 2, ...

T: intervalo, em anos, entre ocorrências de vazões que ultrapassam um certo valor v.

Passamos ao cálculo da probabilidade de que o número de anos seja k.

Notamos que $T = k$ equivale a $X_1 \leq v$, $X_2 \leq v, \ldots, X_{k-1} \leq v$, $X_k > v$.

Supondo que $X_1, X_2, \ldots$ são independentes, calculamos

$$
\begin{aligned}
P(T = k) &= P((X_1 \leq v) \cap (X_2 \leq v) \cap \ldots \cap (X_{k-1} \leq v) \cap (X_k > v)) \\
&= P(X_1 \leq v) \times P(X_2 \leq v) \times \ldots \times P(X_{k-1} \leq v) \times P(X_k > v) \\
&= \{P(X \leq v)\}^{k-1} \times P(X > v), \quad k = 1,2,\ldots
\end{aligned}
$$

Portanto, a distribuição de T é geométrica com parâmetro $p = P(X > v)$.

Período de retorno

O período de retorno (TR) é o tempo médio entre excedências de um valor especificado (v).

$$TR = E(T) = \frac{1}{\mathrm{P}(X > v)}.$$

TR é o recíproco da probabilidade de excedência.