# Capítulo 5 – Distribuições contínuas

## 5.1 - Distribuição uniforme

Embora exista uma distribuição com o mesmo nome para v.a. discretas tratam-se de 2 **distribuições diferentes**

**Definição 5.10 – Distribuição uniforme**

Diz-se que uma variável aleatória contínua $X$ tem distribuição uniforme no intervalo $(\alpha, \beta)$, com $\alpha < \beta$, quando a função densidade é da forma,

$$f(x|\alpha,\beta) = \begin{cases} \dfrac{1}{\beta - \alpha} & (\alpha < x < \beta) \\ 0 & (outros\ x) \end{cases}$$

Simbolicamente, $X \sim U(\alpha, \beta)$

# Capítulo 5 – Distribuições contínuas

- **Função de distribuição:**

$$F(x|\alpha,\beta) = \begin{cases} 0 & x \le \alpha \\ \dfrac{x-\alpha}{\beta-\alpha} & \alpha < x < \beta \\ 1 & x \ge \beta \end{cases}$$

$$P(X \le x) = \frac{x-\alpha}{\beta-\alpha} = \frac{1}{\beta-\alpha}(x-\alpha)$$


- **Expressão geral dos momentos em relação à origem**

$$\boldsymbol{E(X^k) = \int_{\alpha}^{\beta} x^k \frac{1}{\beta-\alpha} dx = \frac{1}{\beta-\alpha}\left[\frac{x^{k+1}}{k+1}\right]_{\alpha}^{\beta} = \frac{\beta^{k+1}-\alpha^{k+1}}{(\beta-\alpha)(k+1)}}$$

# Capítulo 5 – Distribuições contínuas

$$\boldsymbol{E(X) = \frac{\alpha + \beta}{2}};\ \ \boldsymbol{Var(X) = \frac{(\beta - \alpha)^2}{12}};\quad \boldsymbol{\gamma_1 = 0}$$

**Uniforme (0, 1)** O caso particular em que, $\alpha = 0;\ \beta = 1$, isto é, $X \sim U(0,1)$, é o de maior interesse

$$F(x|\alpha,\beta) = \begin{cases} 0 & x \leq 0 \\ x & 0 < x < 1 \\ 1 & x \geq 1 \end{cases}$$

$$E\left(X^k\right) = \frac{1}{k+1};\qquad Var(X) = \frac{1}{12};\qquad \boldsymbol{\gamma_1 = 0}$$

# Capítulo 5 – Distribuições contínuas

- **Teorema 5.1 – Transformação uniformizante**

Em certas condições,

- $X$ v.a. continua com função distribuição $F_X(X)$;

- $F_X(X)$ estritamente crescente no intervalo $(a, b)$ ;

$$F_X(X) = \begin{cases} 0 & x < a \\ F_X(x) & a \leq x < b \\ 1 & x \geq b \end{cases}$$

$a = -\infty$ $ou$ $b = +\infty$ mas não os dois

Seja $F_X(x)$ a função distribuição de uma v.a. $X$. $\quad Y = F_X(X)$

$$F_Y(y) = P(Y \leq y) = P(F_X(X) \leq y) = P\left(X \leq {F_X}^{-1}(y)\right) = F_X\left({F_X}^{-1}(y)\right) = \text{y}$$

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ y & 0 \leq y < 1 \\ 1 & y \geq 1 \end{cases} \quad \Leftrightarrow \; Y = F_X(X) \sim U(0,1)$$

A mudança de variável $Y = F_X(X)$ é chamada **transformação uniformizante** e tem grande aplicação nos problemas de simulação.

# Capítulo 5 – Distribuições contínuas

inversamente se $Y \sim U(0,1)$ então $X = F^{-1}(Y)$ é contínua e tem função distribuição $F_X(x)$.

# Capítulo 5 – Distribuições contínuas

## Teorema 5.1 – Transformação uniformizante

a) Seja $X$ uma v.a. contínua com função de distribuição estritamente crescente no intervalo aberto $(a, b)$ onde $X$ tem densidade positiva [este intervalo pode ser $a = -\infty$ ou $b = +\infty$, tem-se $F_X(a) = 0$ e $F_X(b) = 1$ ]. Então a v.a. $Y = F_X(X) \sim U(0,1)$

b) Seja $Y$ uma variável aleatória com distribuição $U(0,1)$, e $F_X(x)$ a função de distribuição de uma v.a. contínua. Supõe-se que $F_X(x)$ é estritamente crescente no intervalo aberto $(a, b)$ onde $X$ tem densidade positiva [pode ser $a = -\infty$ *ou* $b = +\infty$ *, tem-se* $F_X(a) = 0$ *e* $F_X(b) = 1$ ]. Então a variável aleatória $X = F^{-1}(Y)$ é contínua com função de distribuição$F_X(x)$.

# Capítulo 5 – Distribuições contínuas

## Capítulo 5

**Exerc.10.** A duração de pequenos anúncios (entre 5 e 12 segundos) tem distribuição uniforme.

$$X \sim U\left(\overbrace{5}^{\alpha}, \overbrace{12}^{\beta}\right) \Rightarrow f_X(x) = \frac{1}{5-12} = \frac{1}{7} \quad 5 < x < 12$$

a) Determine a função distribuição.

$$F_X(x) = \int_5^x f_X(x)dx = \int_5^x \frac{1}{7}dx = \frac{x}{7}\Big]_5^x = \begin{cases} 0 & x < 5 \\ \dfrac{x-5}{7} & 5 \le x < 7 \\ 1 & x \ge 7 \end{cases}$$

b) Calcule a probabilidade de um pequeno anúncio ter duração inferior a 8 segundos.

$$P(X < 8) = F_X(8) = \frac{8-5}{7} = \frac{3}{7}$$

c) Calcule e interprete : $P(X > 6|X \le 10)$

$$P(X > 6|X \le 10) = \frac{P(6 < X \le 10)}{P(X \le 10)} = \frac{F_X(10) - F_X(6)}{F_X(10)} = \frac{\frac{5}{7} - \frac{1}{7}}{\frac{10}{7}} = \frac{4}{10}$$

# Capítulo 5 – Distribuições contínuas

## Capítulo 5 – Exercício 10

**d)** Calcule a média e desvio padrão da duração de pequenos anúncios .

$$E(X) = \frac{\alpha + \beta}{2} = \frac{5 + 12}{2} = \frac{17}{2}; \quad Var(X) = \frac{(\beta - \alpha)^2}{12} = \frac{\mathbf{(12 - 5)^2}}{\mathbf{12}} \quad \sigma_X = \sqrt{\frac{\mathbf{49}}{\mathbf{12}}}$$

**Exerc.12.** Seja $X$ uma v.a. contínua com função densidade

$$f_X(x) = \frac{x}{2} \quad (0 < x < 2)$$

Mostre que $Y = F_X(X)$ tem distribuição uniforme.

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \int_0^x \frac{x}{2} dx = \frac{x^2}{4} & 0 \leq x < 2 \\ 1 & x \geq 2 \end{cases} \qquad F_Y(y) = \begin{cases} 0 & y < 0 \\ y & 0 \leq y < 1 \\ 1 & y \geq 1 \end{cases}$$

$$\Rightarrow \quad F_Y(y) = P(Y \leq y) = P(F_X(X) \leq y) = P\left(\frac{X^2}{4} \leq y\right) = P(X^2 \leq 4y)$$

$$= P\left(X \leq \sqrt{4y}\right) = \frac{\left(\sqrt{4y}\right)^2}{4} = \text{y}$$

# Capítulo 5 – Distribuições contínuas

Ex1. O presidente da câmara de Lisboa pretende realizar um projecto de recuperação num dos bairros da cidade. Estima-se que o custo (em milhares de €) associado ao projecto é de 1500 m€. Contudo sabe-se também que este tipo de projectos sofrem derrapagens orçamentais (em milhares de €) que seguem uma distribuição normal de média  20 e variância 25). Dadas as restrições orçamentais a câmara tem um plafond de 3000 m.€ para realizar o projecto. Acha que o projecto deve ser implementado?

Ex 2. Gestores Financeiros, Lda., compra e vende um grande número de acções para as várias contas que gere. Uma parte desta carteira é composta por 10 acções da empresa A e 8 acções da empresa B. Os preços, em €, de A e B seguem uma distribuição normal, respectivamente, com médias e variâncias 10, 16 e 12, 9. Qual a média e a variância do valor da carteira. Como gerente de portfólio da empresa é-lhe pedida opinião sobre o intervalo de valor da carteira com uma probabilidade de 90%?

# Capítulo 5 – Distribuições contínuas

## 5.2 - Distribuição normal

**Definição 5.11 – Distribuição normal**

Uma variável aleatória $X$ tem distribuição normal com parâmetros $\mu$ e $\sigma^2$ quando a função densidade é da forma

$$f(x|\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

$$-\infty < x < +\infty; \quad -\infty < \mu < +\infty; \quad 0 < \sigma^2 < +\infty$$

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

# Capítulo 5 – Distribuições contínuas

## 5.2 - Distribuição normal

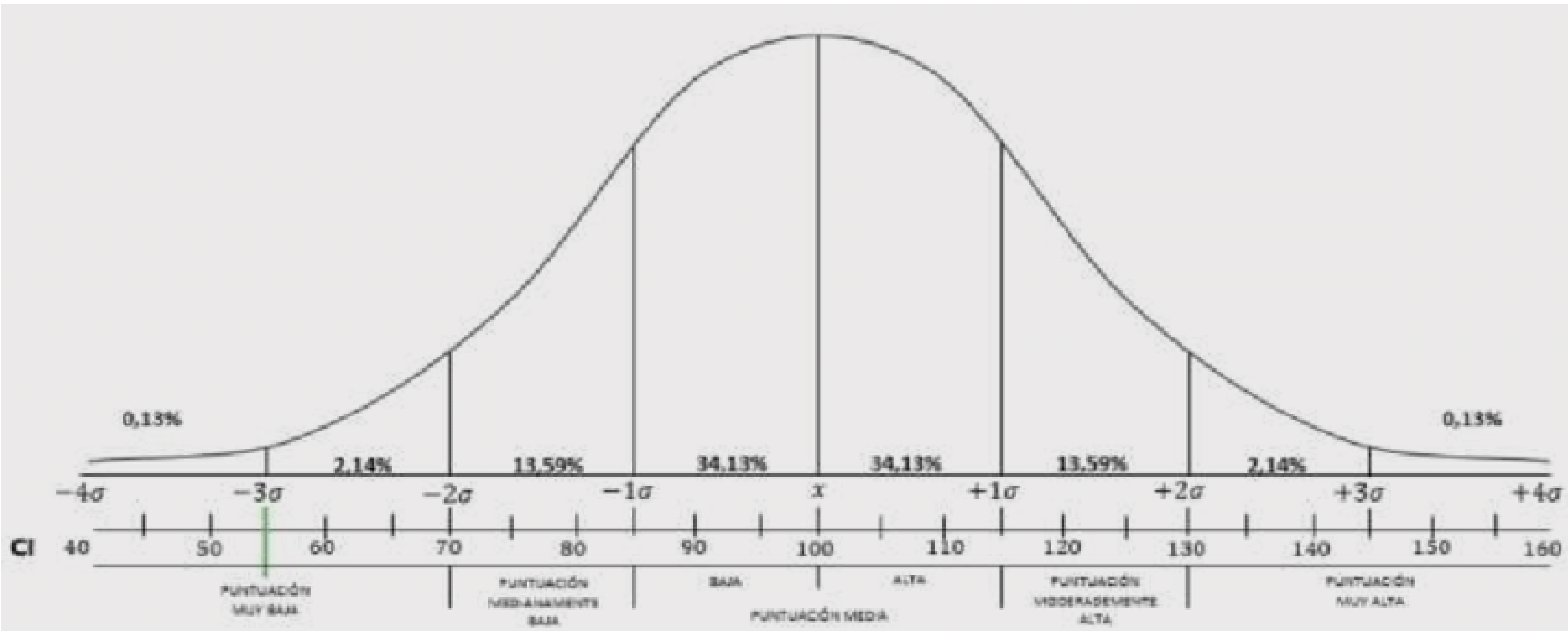


**Fig. 5.1a –** Funções densidade da distribuição normal com a mesma variância $\sigma^2 = 1$

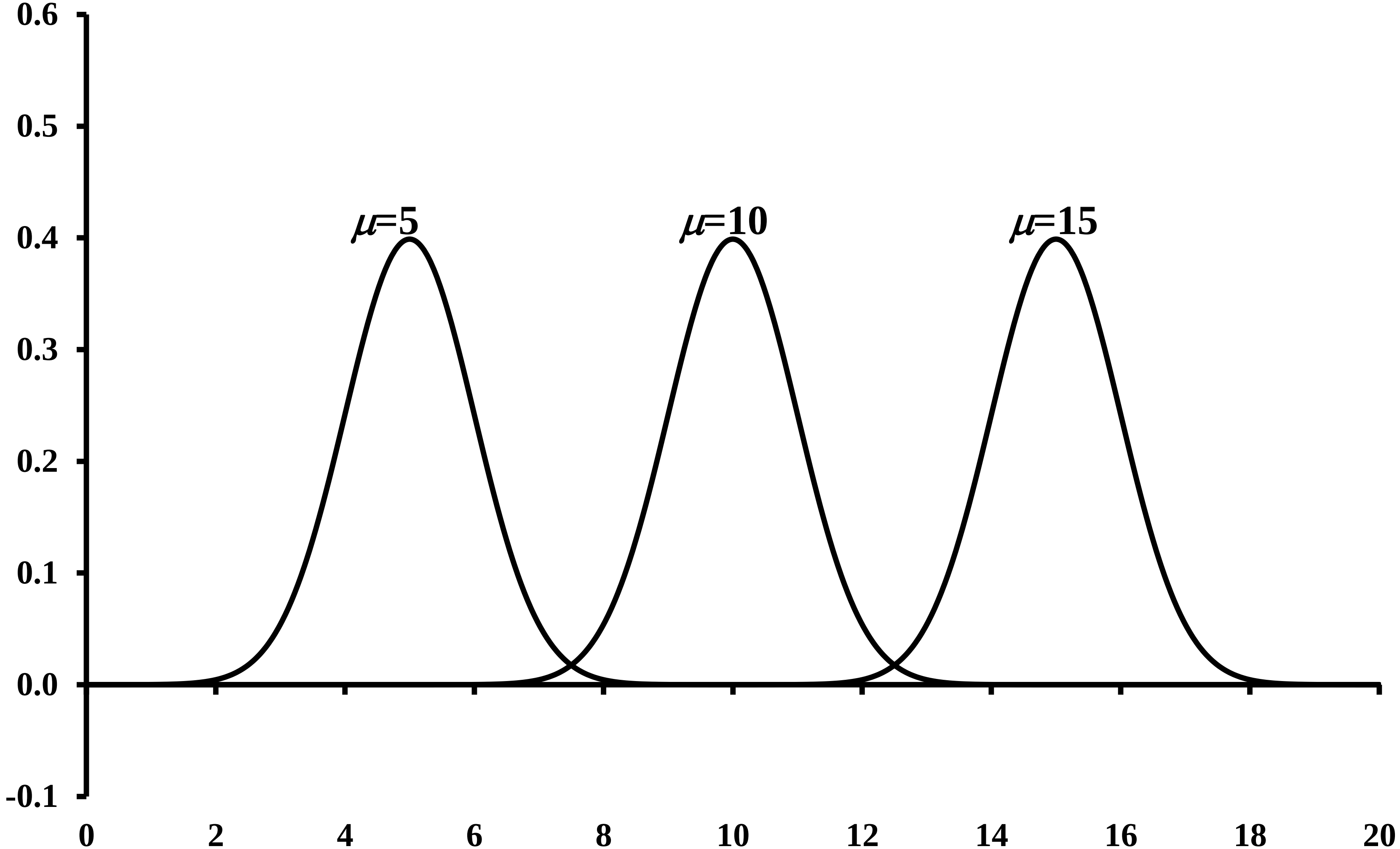
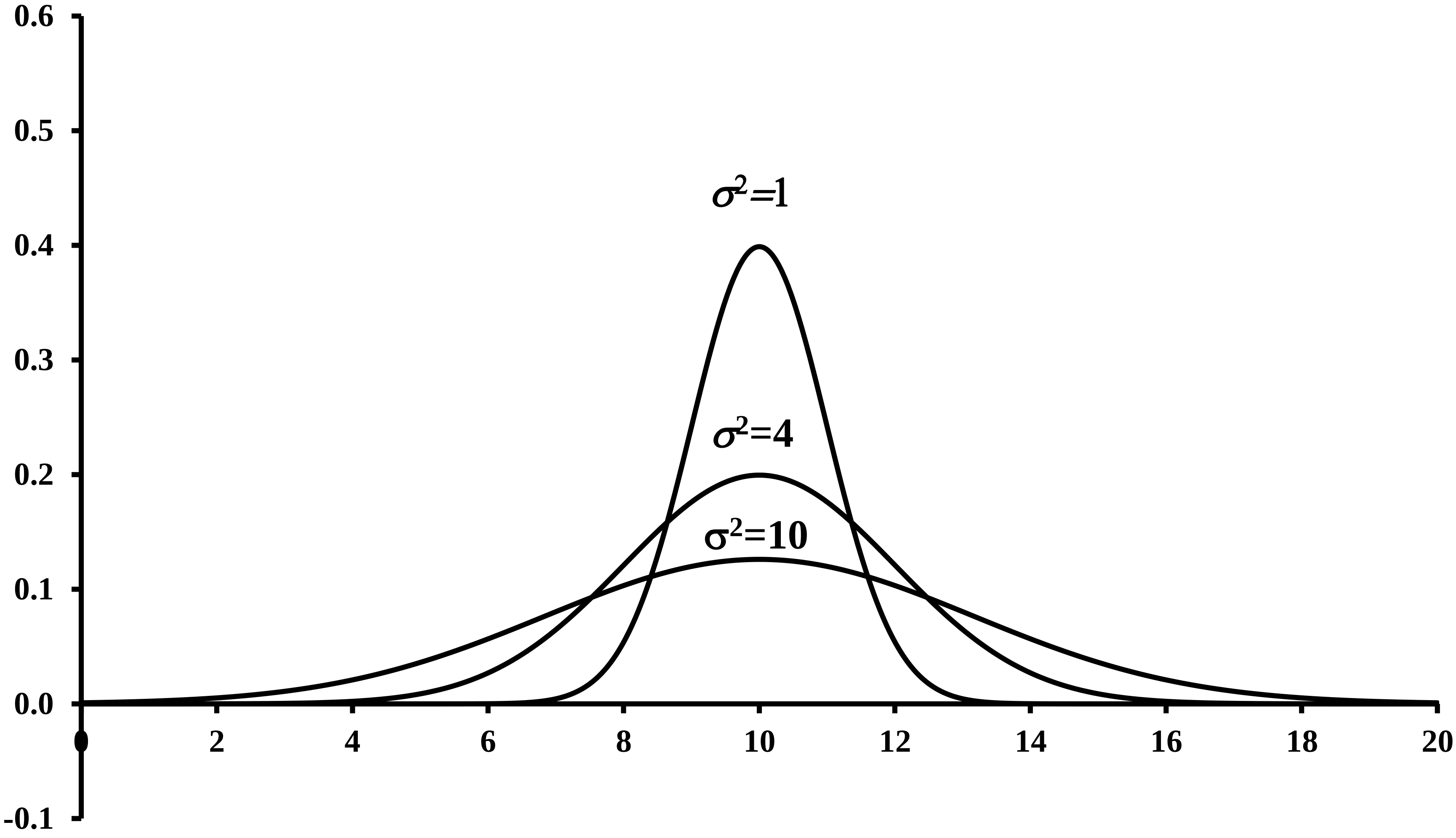


**Fig. 5.1b –** Funções densidade da distribuição normal com a mesma média $\mu = 10$



# Capítulo 5 – Distribuições contínuas

- **Função de distribuição**

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

para o qual **não se conhece solução analítica**. Os valores da função de distribuição têm então de ser calculados utilizando métodos de **análise numérica**.

**Distribuição normal estandardizada** caso particular, $\mu = 0$ e $\sigma^2 = 1$

Processo de estandardização de uma v.a. $X \sim N(\mu, \sigma^2) \rightarrow Z \sim N(0, 1)$

Seja $Y = X - \mu \Rightarrow \mathrm{E}(Y) = \mathrm{E}(X-\mu) = \mathrm{E}(X) - \mu = \mu - \mu = 0$

Seja $Y = \frac{X-\mu}{\sigma} \Rightarrow Var\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2} * Var(Y) = \frac{\sigma^2}{\sigma^2} = 1$

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

# Capítulo 5 – Distribuições contínuas

- **Distribuição normal estandardizada** $Z \sim N(0,1)$

  - função densidade : $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$

  - função distribuição: $\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$

Nota importante: só existe tabela para a variável $Z \sim N(0,1)$

A tabela 4 refere-se à função $\Phi(z) = \mathrm{P}(Z \leq z)$

A tabela 5 diz respeito à função inversa $\Phi^{-1}(z)$

$\Phi^{-1}(z)$ permite determinar, para certos valores de $\Phi(z)$, a respectiva abcissa $z$.

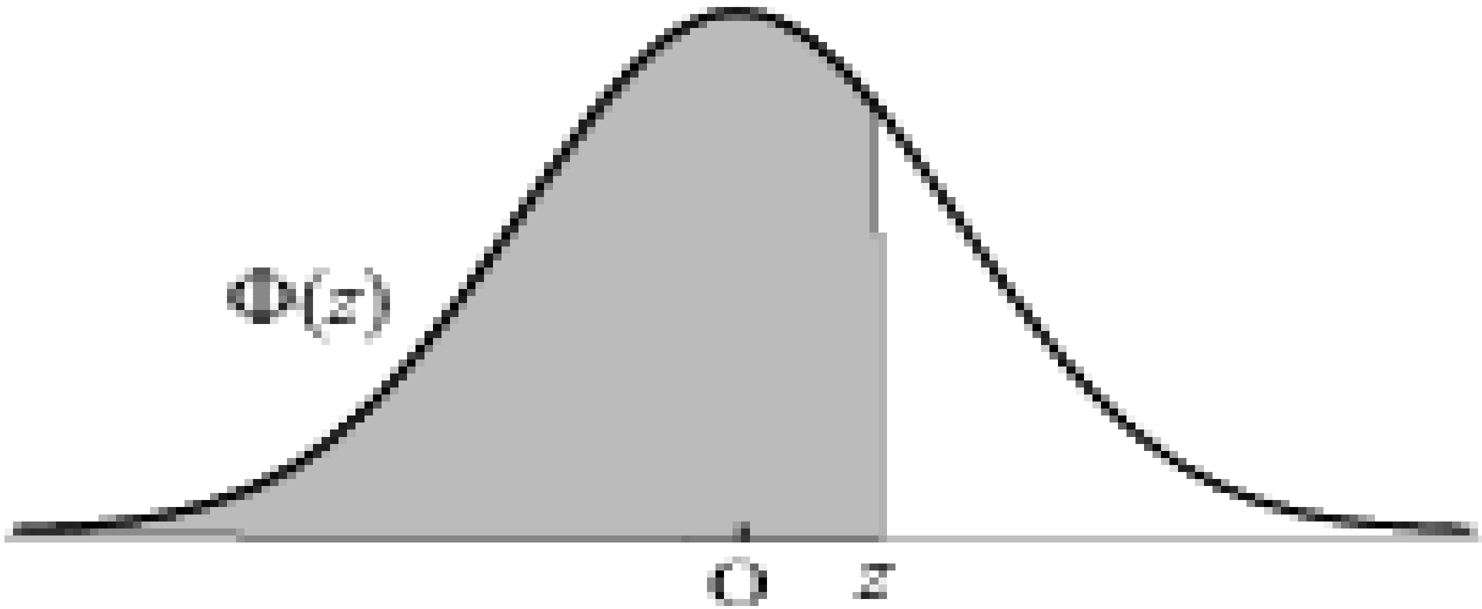
$z\colon P(Z > z) = \alpha \quad \Phi^{-1}(z) = \alpha$

# Capítulo 5 – Distribuições contínuas

TABELA 4 – DISTRIBUIÇÃO NORMAL STANDARDIZADA

Função de distribuição

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$



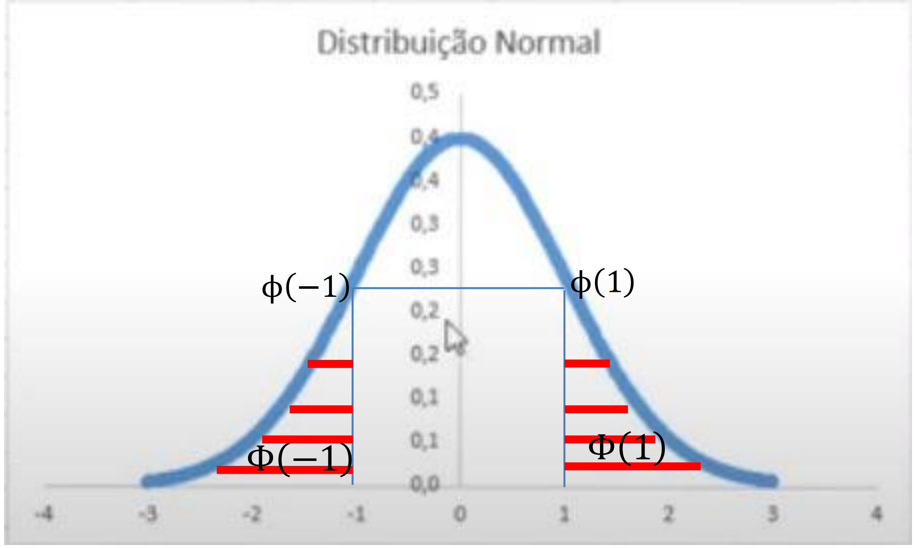
# Capítulo 5 – Distribuições contínuas

- A função densidade normal é simétrica em relação à recta $x = \mu$

$$f(x-\mu) = f(x+\mu)$$

Em particular, tratando-se da distribuição normal estandardizada,

$$\phi(x) = \phi(-x) \qquad \Phi(-x) = 1 - \Phi(x)$$ Importante



# TABELA 4 – DISTRIBUIÇÃO NORMAL – Função de distribuição

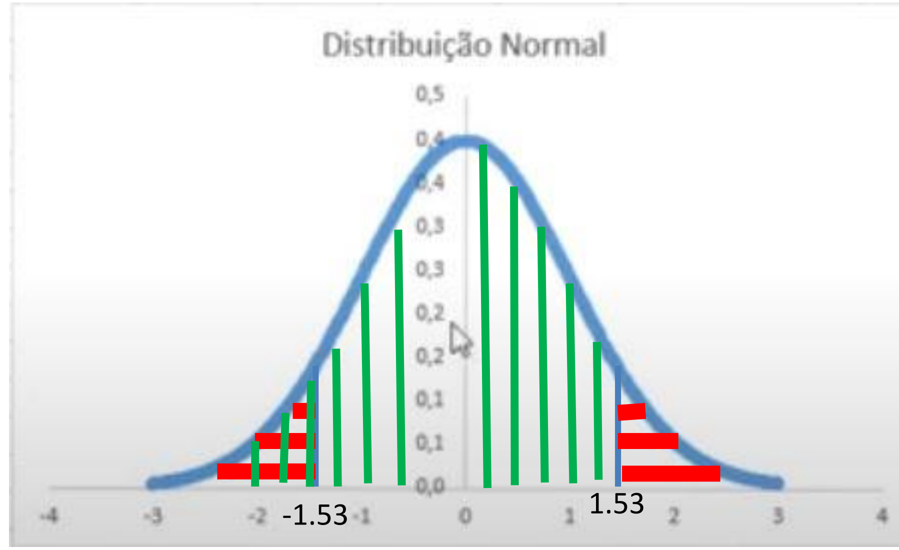
| z | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|---|---|---|---|---|---|---|---|---|---|---|
| .0 | .5000 | .5040 | .5080 | .5120 | .5160 | .5199 | .5239 | .5279 | .5319 | .5359 |
| .1 | .5398 | .5438 | .5478 | .5517 | .5557 | .5596 | .5636 | .5675 | .5714 | .5753 |
| .2 | .5793 | .5832 | .5871 | .5910 | .5948 | .5987 | .6026 | .6064 | .6103 | .6141 |
| .3 | .6179 | .6217 | .6255 | .6293 | .6331 | .6368 | .6406 | .6443 | .6480 | .6517 |
| .4 | .6554 | .6591 | .6628 | .6664 | .6700 | .6736 | .6772 | .6808 | .6844 | .6879 |
| .5 | .6915 | .6950 | .6985 | .7019 | .7054 | .7088 | .7123 | .7157 | .7190 | .7224 |
| .6 | .7257 | .7291 | .7324 | .7357 | .7389 | .7422 | .7454 | .7486 | .7517 | .7549 |
| .7 | .7580 | .7611 | .7642 | .7673 | .7704 | .7734 | .7764 | .7794 | .7823 | .7852 |
| .8 | .7881 | .7910 | .7939 | .7967 | .7995 | .8023 | .8051 | .8078 | .8106 | .8133 |
| .9 | .8159 | .8186 | .8212 | .8238 | .8264 | .8289 | .8315 | .8340 | .8365 | .8389 |
| 2.5 | .9938 | .9940 | .9941 | .9943 | .9945 | .9946 | .9948 | .9949 | .9951 | .9952 |
| 2.6 | .9953 | .9955 | .9956 | .9957 | .9959 | .9960 | .9961 | .9962 | .9963 | .9964 |
| 2.7 | .9965 | .9966 | .9967 | .9968 | .9969 | .9970 | .9971 | .9972 | .9973 | .9974 |
| 2.8 | .9974 | .9975 | .9976 | .9977 | .9977 | .9978 | .9979 | .9979 | .9980 | .9981 |
| 2.9 | .9981 | .9982 | .9982 | .9983 | .9984 | .9984 | .9985 | .9985 | .9986 | .9986 |
| 3.0 | .9987 | .9987 | .9987 | .9988 | .9988 | .9989 | .9989 | .9989 | .9990 | .9990 |
| 3.1 | .9990 | .9991 | .9991 | .9991 | .9992 | .9992 | .9992 | .9992 | .9993 | .9993 |
| 3.2 | .9993 | .9993 | .9994 | .9994 | .9994 | .9994 | .9994 | .9995 | .9995 | .9995 |
| 3.3 | .9995 | .9995 | .9995 | .9996 | .9996 | .9996 | .9996 | .9996 | .9996 | .9997 |
| 3.4 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9997 | .9998 |

# Capítulo 5 – Distribuições contínuas

$$P(Z < 2.53) = 0.9943$$

$$P(Z > 1.53) = 1 - P(Z < 1.53)$$

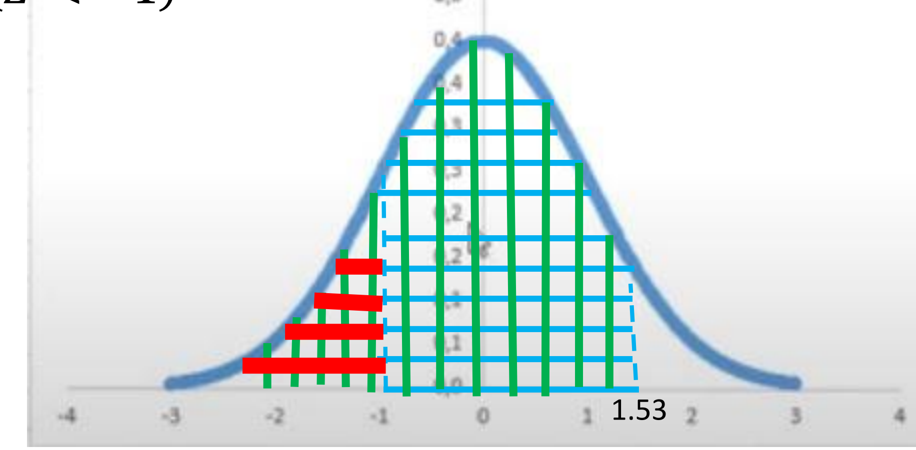
$$P(Z < -1.53) = 1 - P(Z < 1.53)$$




$$P(-1 \leq Z < 1.53) = P(Z < 1.53) - P(Z < -1)$$

$$= P(Z < 1.53) - [1 - P(Z < 1)]$$

$$= P(Z < 1{,}53) + P(Z < 1) - 1$$



# Capítulo 5 – Distribuições contínuas

## TABELA 5 – DISTRIBUIÇÃO NORMAL: $\phi^{-1}(z)$

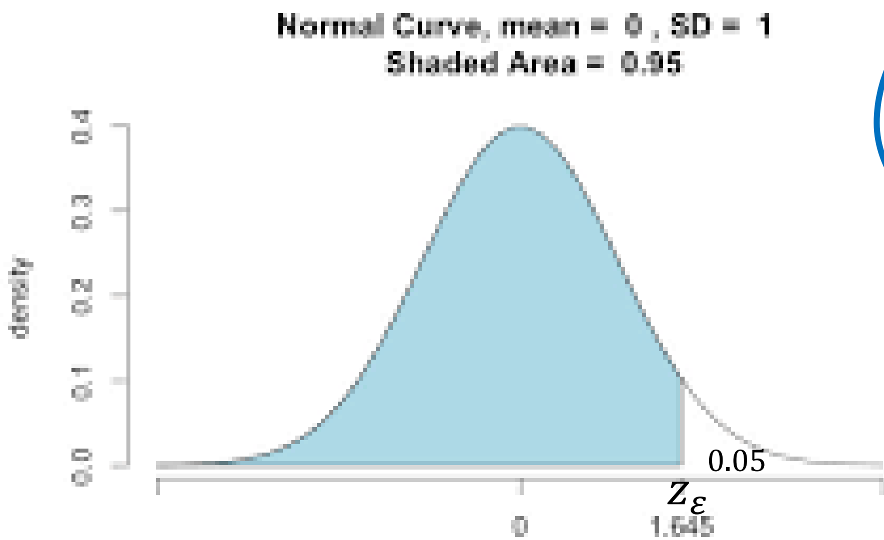
| $\varepsilon$ | .0005 | .0010 | .0050 | .0100 | .0200 | .0250 | .0500 | .1000 | .2000 | .3000 | .4000 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $Z_\varepsilon$ | 3.290 | 3.090 | 2.576 | 2.326 | 2.054 | 1.960 | 1.645 | 1.282 | 0.842 | 0.524 | 0.253 |
| $Z_{\varepsilon/2}$ | 3.481 | 3.290 | 2.807 | 2.576 | 2.326 | 2.241 | 1.960 | 1.645 | 1.282 | 1.036 | 0.842 |

- $z_\varepsilon : P(Z > z_\varepsilon) = \epsilon;\ z_{\varepsilon/2} : P\left(|Z| > z_{\varepsilon/2}\right) = \varepsilon$

Esta tabela usa-se quando há necessidade de calcular certos intervalos de $X$ associados a probabilidades que são dadas:

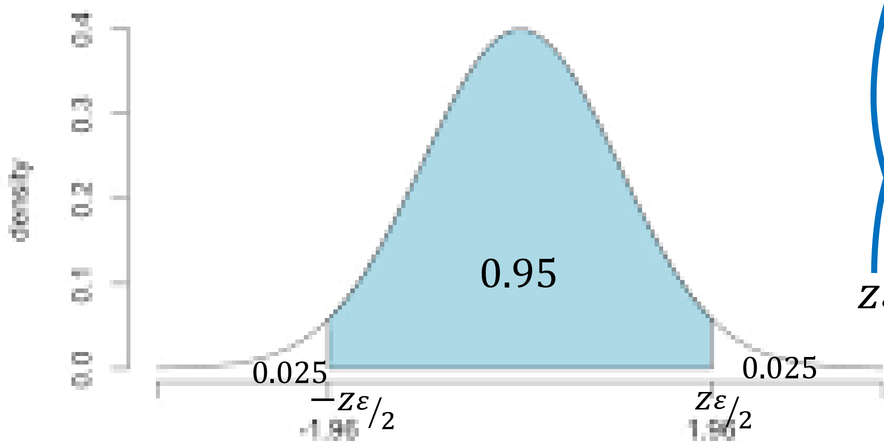
- Calcular $k$ tal que $P(|Z| > k) = 0{,}05$

# Capítulo 5 – Distribuições contínuas




| $\varepsilon$ | .0005 | .0010 | .0050 | .0100 | .0200 | .0250 | .0500 | .1000 | .2000 | .3000 | .4000 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $Z_\varepsilon$ | 3.290 | 3.090 | 2.576 | 2.326 | 2.054 | 1.960 | 1.645 | 1.282 | 0.842 | 0.524 | 0.253 |
| $Z_{\varepsilon/2}$ | 3.481 | 3.290 | 2.807 | 2.576 | 2.326 | 2.241 | 1.960 | 1.645 | 1.282 | 1.036 | 0.842 |

$$z_\varepsilon = ? : P(Z > z_\varepsilon) = 0.05$$




| $\varepsilon$ | .0005 | .0010 | .0050 | .0100 | .0200 | .0250 | .0500 | .1000 | .2000 | .3000 | .4000 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $Z_\varepsilon$ | 3.290 | 3.090 | 2.576 | 2.326 | 2.054 | 1.960 | 1.645 | 1.282 | 0.842 | 0.524 | 0.253 |
| $Z_{\varepsilon/2}$ | 3.481 | 3.290 | 2.807 | 2.576 | 2.326 | 2.241 | 1.960 | 1.645 | 1.282 | 1.036 | 0.842 |

$$z_{\varepsilon/2} = ? : P\left(-z_{\varepsilon/2} < Z \leq z_{\varepsilon/2}\right) = 0.95$$

$$z_{\varepsilon/2} = ? : P\left(Z \leq -z_{\varepsilon/2}\right) + P\left(Z > z_{\varepsilon/2}\right) = 0.05$$

$$z_\varepsilon = ? : P(Z > z_\varepsilon) = 0.025$$

# Capítulo 5 – Distribuições contínuas

- **Exemplo 5.12 –** A resistência à compressão de amostras de cimento de um certo tipo é uma variável aleatória que pode ser modelada por uma distribuição normal com média 6000 kg/cm$^2$ e desvio padrão 100 kg/cm$^2$. Calcule: $X \sim N(6000, 100^2)$

a) Probabilidade de uma amostra cimento ter uma resistência superior a 6150 kg/cm$^2$.

$$P(X > 6150) = 1 - P(X \leq 6150)$$
$$= 1 - normalcdf\left(\overbrace{-10000}^{lim.inf.}, \overbrace{6150}^{lim.sup.}, 6000, 100\right) = 1 - 0.9332$$

ou

$$P(X > 6150) = 1 - P(X \leq 6150) = 1 - P\left(\frac{X - \mu}{\sigma} \leq \frac{6150 - 6000}{100}\right)$$
$$= 1 - P(Z \leq 1.5) = 1 - \Phi(1.5) = 1 - 0.9932$$

# Capítulo 5 – Distribuições contínuas

b) Probabilidade de uma amostra de cimento ter uma resistência entre 5900 kg/cm² e 5950 kg/cm².

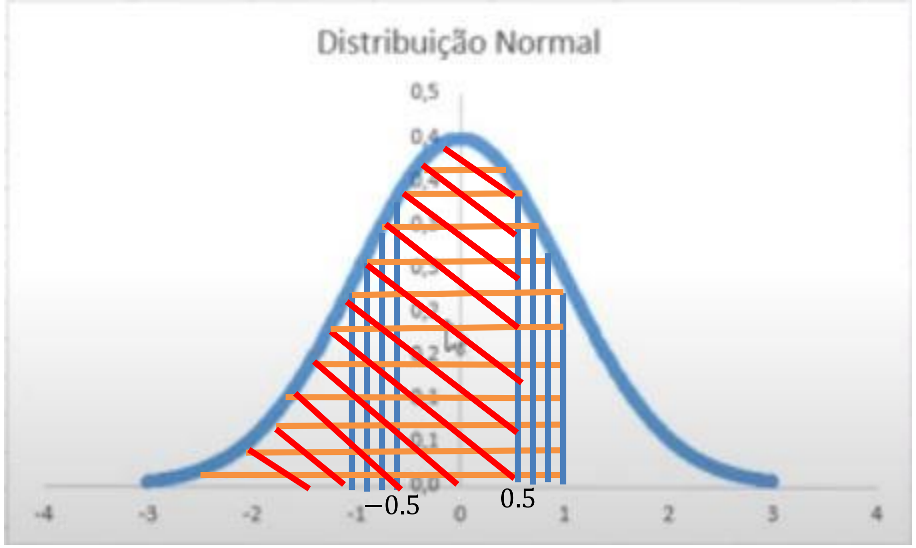
$$P(5900 < X < 5950) = normalcdf\left(\overbrace{5900}^{lim.inf.}, \overbrace{5950}^{lim.sup.}, 6000, 100\right) = 0.1499$$

ou

$$P\left(\frac{5900-6000}{100} < \frac{X-\mu}{\sigma} \leq \frac{5950-6000}{100}\right) = P(-1 < Z \leq -0.5)$$

$$= \Phi(-0.5) - \Phi(-1) = \Phi(1) - \Phi(0.5)$$

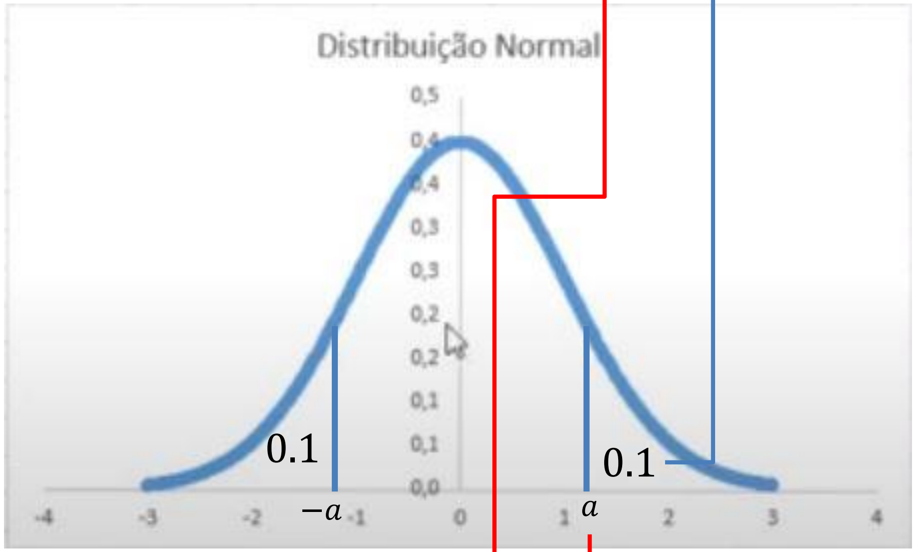
$$= 0.8413 - 0.6915 = 0.1499$$



# Capítulo 5 – Distribuições contínuas

| $\varepsilon$ | .0005 | .0010 | .0050 | .0100 | .0200 | .0250 | .0500 | .1000 | .2000 | .3000 | .4000 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $Z_\varepsilon$ | 3.290 | 3.090 | 2.576 | 2.326 | 2.054 | 1.960 | 1.645 | 1.282 | 0.842 | 0.524 | 0.253 |
| $Z_{\varepsilon/2}$ | 3.481 | 3.290 | 2.807 | 2.576 | 2.326 | 2.241 | 1.960 | 1.645 | 1.282 | 1.036 | 0.842 |

c) A resistência *R* que é excedida por 90% das amostras de cimento.




$$r = ?: P(R > r) = 0.9$$

$$r = ?: P(R > r) = 1 - P(R \leq r) = 0.9$$

$$P(R \leq r) = 0.1$$

$$\Rightarrow r = invnorm(0.1, 6000, 100)$$
$$= 5871.845$$

ou

$$P(R > r) = 0.9 \Leftrightarrow P\left(\frac{R-\mu}{\sigma} \leq \underbrace{\frac{r-6000}{100}}_{-a}\right) = P(Z > a) = 0.1$$

$$\Rightarrow \frac{r-6000}{100} = -1.282 \Leftrightarrow r = 5871.845$$

# Capítulo 5 – Distribuições contínuas

**Teorema 5.2 –** Se as variáveis aleatórias $X_i,\ i = 1, 2, \cdots, k$ são independentes e $\alpha_i$ constantes, então,

$$X_i \sim N(\mu_i, \sigma_i^2) \Rightarrow Y = \sum_{i=1}^{k} \alpha_i X_i \sim N\left(\sum_{i=1}^{k} \alpha_i \mu_i, \sum_{i=1}^{k} \alpha_i^2 \sigma_i^2\right)$$

**Corolário 5.1 –** Se as variáveis aleatórias $X_i,\ i = 1, 2, \cdots, k$ são independentes e identicamente distribuídas então,

$$X_i \sim N(\mu, \sigma^2) \Rightarrow Y = \sum_{i=1}^{k} X_i \sim N(k\mu, k\sigma^2)$$

# Capítulo 5 – Distribuições contínuas

Ex1. O presidente da câmara de Lisboa pretende realizar um projecto de recuperação num dos bairros da cidade. Estima-se que o custo (em milhares de €) associado ao projecto é de 1500 m€. Contudo sabe-se também que este tipo de projectos sofrem derrapagens orçamentais (em milhares de €) que seguem uma distribuição normal de média  200 e variância 25). Dadas as restrições orçamentais a câmara tem um plafond de 3000 m.€ para realizar o projecto. Acha que o projecto deve ser implementado?

$X$- derrapagem orçamental em $10^3€ \sim N(200, 25)$

$C$- custo projecto$= 1500 + X$ $\quad P(C > 3000) = P(1500 + X > 3000) = P(X > 1500)$

$$= normcdf\left(\overbrace{1500}^{lim.inf}, \overbrace{10000}^{lim.sup}, 200, 5\right) \approx 0$$

ou

$$P(C > 3000) = P(1500 + X > 3000) = P(X > 1500) = 1 - P(X \leq 1500)$$
$$= 1 - P\left(\frac{X - \mu}{\sigma} \leq \frac{1500 - 200}{5}\right) = 1 - P(Z \leq 260) = 1 - \Phi(260) \approx 1 - 1 = 0$$

# Capítulo 5 – Distribuições contínuas

Ex 2. Gestores Financeiros, Lda., compra e vende um grande número de acções para as várias contas que gere. Uma parte desta carteira é composta por 10 acções da empresa A e 8 acções da empresa B. Os preços, em €, de A e B seguem uma distribuição normal, respectivamente, com médias e variâncias 10, 16 e 12, 9. Qual a média e a variância do valor da carteira. Como gerente de portfólio da empresa é-lhe pedida opinião sobre o intervalo de valor da carteira com uma probabilidade de 90%?

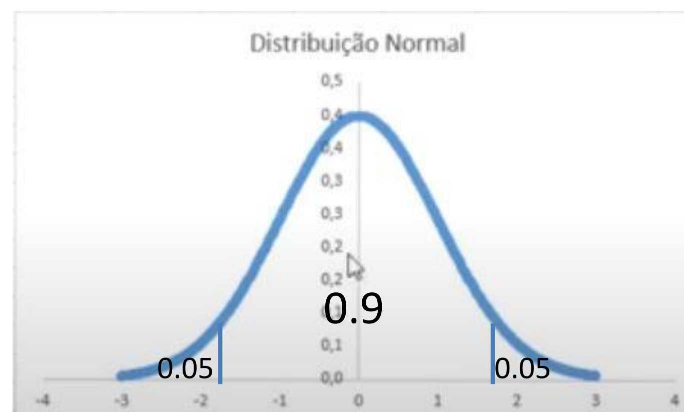
$X_i$- preço das acções empresa A$\sim N(10, 16)$ $Y_j$- preço das acções empresa B$\sim N(12, 9)$

$$\sum_{i=1}^{10} X_i \sim N(10 * 10, 10 * 16) \quad \sum_{j=1}^{8} Y_i \sim N(8 * 12, 8 * 9)$$

$$V = \sum_{i=1}^{10} X_i + \sum_{j=1}^{8} Y_i \Rightarrow V \sim N(100 + 96, 160 + 72)$$

$$V = ? : P(v_1 < V < v_2) = 0.9$$

$$v_1 : P(V < v_1) = 0.05 \quad v_2 : P(V > v_2) = 0.05$$



# Capítulo 5 – Distribuições contínuas

$$V = \sum_{i=1}^{10} X_i + \sum_{j=1}^{8} Y_i \Rightarrow V \sim N(100 + 96, 160 + 72)$$

$$V = ?: P(v_1 < V < v_2) = 0.9$$

$$v_1: P(V < v_1) = 0.05 \Rightarrow v_1 = invnorm\left(0.05, 196, \sqrt{232}\right) = 170.95$$

$$v_2: P(V > v_2) \Rightarrow v_2 = invnorm\left(0.95, 196, \sqrt{232}\right) = 221.05$$

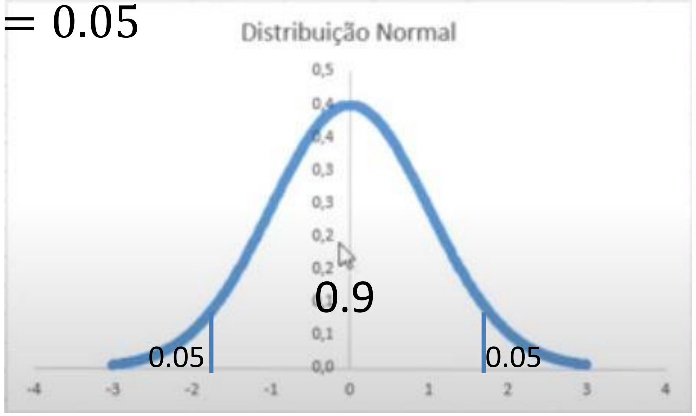
$ou$

$$v_1: P(V < v_1) = 0.05 \Leftrightarrow P(V < v_1) = P(V > v_2) = 0.05$$

$$P\left(\frac{V-\mu}{\sigma} > \frac{v_2 - 196}{15.23}\right) = 0.05 \Leftrightarrow P\left(Z > \frac{v_2 - 196}{15.23}\right) = 0.05$$

$$\frac{v_2 - 196}{15.23} = 1.645 \Leftrightarrow v_2 = 221.06$$

$$\frac{v_1 - 196}{15.23} = -1.645 \Leftrightarrow v_1 = 170.95$$



# Capítulo 5 – Distribuições contínuas

**Corolário 5.2 –** Se as variáveis aleatórias , $X_i,\ i = 1, 2, \cdots, k$ são independentes e identicamente distribuídas,

$$X_i \sim N(\mu, \sigma^2) \Rightarrow \bar{X} = \frac{\sum_{i=1}^{k} X_i}{k} \sim N\left(\mu, \frac{\sigma^2}{k}\right)$$

# Capítulo 5 – Distribuições contínuas

**Teorema 5.3** – Se as variáveis aleatórias $X_i$, $i = 1, 2, \cdots, k$ são normais tais que

$$E(X_i) = \mu_i; \quad Var(X_i) = \sigma_i^2; Cov\left(X_i, X_j\right) = \sigma_{i,j} \quad (i \neq j)$$

Então

$$Y = \sum_{i=1}^{k} \alpha_i X_i \sim N(\mu_Y, \sigma_Y^2)$$

Com

$$\mu_y = \sum_{i=1}^{k} \alpha_i \mu_i$$

e $\sigma_Y^2 = \alpha_1^2\sigma_1^2 + 2\alpha_1\alpha_2\sigma_{12} + 2\alpha_1\alpha_3\sigma_{13} + \cdots + 2\alpha_1\alpha_k\sigma_{1k} +$
$+\alpha_2^2\sigma_2^2 + 2\alpha_2\alpha_1\sigma_{21} + 2\alpha_2\alpha_3\sigma_{23} + \cdots + 2\alpha_2\alpha_k\sigma_{2k} +$
$+ \cdots + 2\alpha_k\alpha_1\sigma_{k1} + 2\alpha_k\alpha_3\sigma_{k3} + \cdots + \alpha_k^2\sigma_k^2$

# Capítulo 5 – Distribuições contínuas

## 5.3. Distribuição exponencial

Ex 1. Participações de falhas nos circuitos eléctricos do modelo novo da Audi são raras. Em certa concessionária ocorrem a uma taxa média de 0,1 participações por mês.

Qual é a probabilidade de a concessionária não receber nenhuma participação nos próximos seis meses?

Foi recebida uma participação à 4 meses atrás, qual a probabilidade de que passem mais de 8 meses até à próxima reclamação ?

Ex 2. A empresa Macarrão & Companhia garante que os seus clientes não esperam mais de 3 minutos pelo atendimento. Sabe-se que o número médio de clientes atendidos por minuto segue um processo de Poisson de taxa média igual a 3. Estará a empresa a fazer publicidade enganosa?

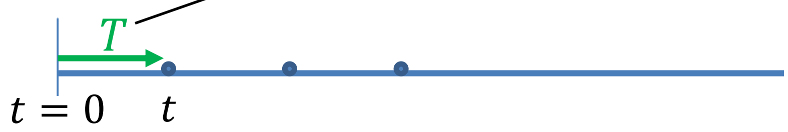
# Capítulo 5 – Distribuições contínuas

## 5.3. Distribuição exponencial

A distribuição exponencial, ou exponencial negativa como também é designada, tem a sua génese associada ao processo de Poisson, muito embora a sua utilização na modelação estatística seja bastante mais ampla;

Seja uma sucessão de eventos que ocorre de acordo com um processo de Poisson de intensidade λ

**tempo de espera** pela chegada da primeira ocorrência




variável aleatória $T$ tem função de distribuição, $F_T(t)$ dada por:

$$F_T(t) = \mathrm{P}(T \leq t) = 1 - \mathrm{P}(T > t) \quad (t > 0)$$

$\mathrm{P}(T > t)$ probabilidade de não haver qualquer chegada no intervalo $(0, t]$

Se se definir a v.a. $X$- nº ocorrências no intervalo $(0, t] \Rightarrow X \sim Po(\lambda)$

$$\mathrm{P}(T > t) = P(X = 0) = e^{-\lambda t} \quad (\lambda > 0, t > 0)$$

# Capítulo 5 – Distribuições contínuas

$$\mathrm{P}(T > t) = 1 - \mathrm{P}(T \leq t) = e^{-\lambda t} \iff \mathrm{P}(T \leq t) = 1 - e^{-\lambda t}$$

$$F_T(t) = \begin{cases} 0 & (t \leq 0) \\ 1 - e^{-\lambda t} & (t > 0) \end{cases}$$

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

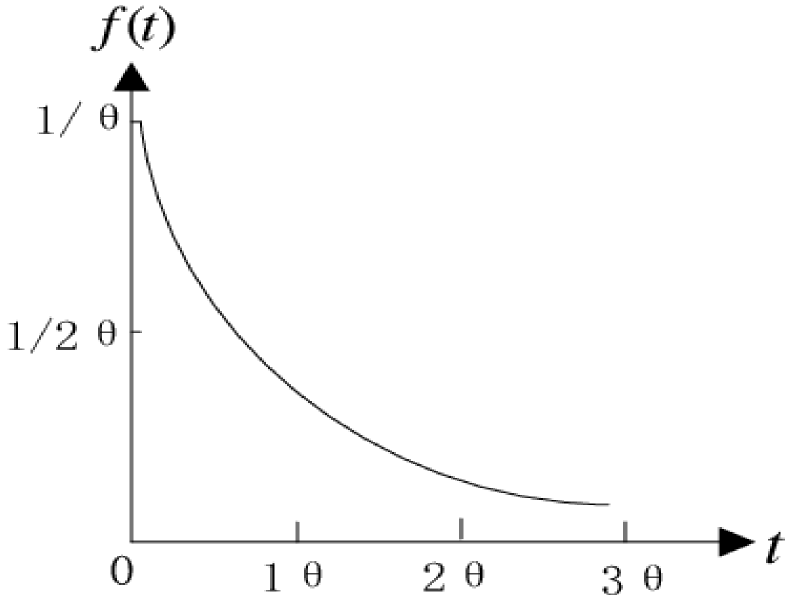
**Definição 5.12 – Distribuição exponencial**

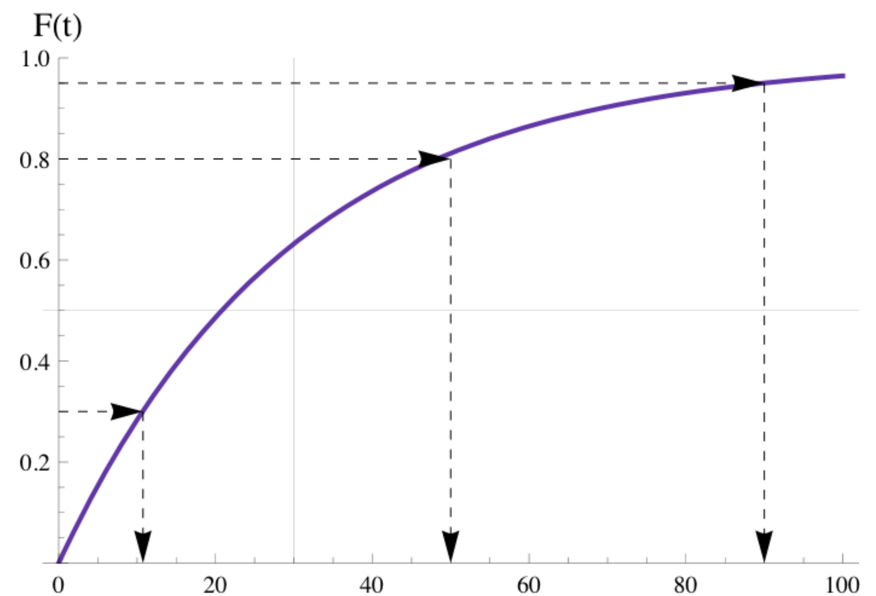
Variável aleatória $X$ com função densidade:

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

Tem distribuição exponencial $X \sim Ex(\lambda)$

# Capítulo 5 – Distribuições contínuas

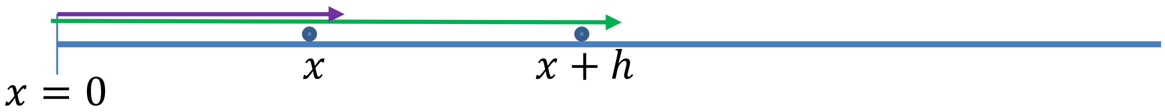



$$\boldsymbol{E(X) = \frac{1}{\lambda}} \; ; \qquad \boldsymbol{Var(X) = \frac{1}{\lambda^2}} \; ;$$

$$\gamma_1 = 2 \qquad \mu_e = ln(2)/\lambda \qquad \mu_e < \mu$$

# Capítulo 5 – Distribuições contínuas




$$\mathrm{P}(X > x) = e^{-\lambda x}$$

$$\boldsymbol{P(X > x + h|X > x)} = \frac{\boldsymbol{P(X > x + h)}}{\boldsymbol{P(X > x)}} = \frac{e^{-\lambda(x+h)}}{e^{-\lambda x}} = e^{-\lambda(x+h)-(-\lambda x)} = e^{-\lambda h}$$

$$\Rightarrow \boldsymbol{P(X > x + h|X > x) = P(X > h)}$$

Falta de memória

# Capítulo 5 – Distribuições contínuas

**Exemplo 5.15** (alterado) **–** Suponha-se que o tempo de vida útil de um dado componente é uma v.a. $X$ com distribuição exponencial de média igual a 600 horas. $X \sim Ex(\lambda) \quad E(X) = \frac{1}{\lambda} = 600 \Leftrightarrow \lambda = 1/600$

**a)** Qual a probabilidade de o componente durar mais de 700 horas?

$$P(X > 700) = e^{-\frac{1}{600} * 700} = e^{-\frac{7}{6}}$$

**b)** Se o componente já tiver durado 400 horas qual a probabilidade de durar ainda mais 700 horas?

$$P(X > 700 | X > 400) = \frac{P(X > 700)}{P(X > 400)} = \frac{e^{-\frac{1}{600} * 700}}{e^{-\frac{1}{600} * 400}} = e^{-\frac{1}{600}(700-400)}$$

$$= e^{-\frac{1}{600} * 300} = e^{-\frac{3}{6}}$$

# Capítulo 5 – Distribuições contínuas

Exerc. Nº29 O número de partículas radioactivas observadas num intervalo de 10 minutos segue um processo de Poisson. $X \sim Po(\lambda)$

a) Supondo que a probabilidade de não observar qualquer partícula num intervalo de 10 minutos é 0,15, determine o ritmo do processo, i.é, em média quanto tempo medeia entre duas observações consecutivas.

$$X \sim Po(\lambda) \Rightarrow P(X=0) = \frac{e^{-\lambda} * \lambda^0}{0!} = e^{-\lambda} = 0{,}15 \quad \Leftrightarrow \quad ln\left(e^{-\lambda}\right) = \ln(0.15)$$

$$\Leftrightarrow -\lambda = -1{,}8971 = \Leftrightarrow \lambda = 1{,}8971 = E(X)$$

Número médio de partículas observadas num minuto $= 1{,}8971/10 = 0{,}1897$

O tempo médio, em minutos, que medeia entre partículas será $= 1/0{,}1897 = 5{,}2715$

$Y - tempo\ entre\ observação\ de\ 2\ particulas\ consecutivas \qquad Y \sim Ex(\lambda^*)$

$$E(Y) = \frac{1}{\lambda^*} \Leftrightarrow 5{,}2715 = \frac{1}{\lambda^*} \Leftrightarrow \lambda^* = \frac{1}{5{,}2715} = 0.1897$$

## 5.4 Distribuição gama. Distribuição do qui-quadrado

A distribuição gama também pode ser introduzida a partir do tempo de espera associado a um processo de Poisson: **Tempo (total) de espera pela *n*-ésima chegada**.

**Função gama** ou factorial generalizado:

$$\Gamma(\alpha) = \int_0^{+\infty} e^{-x} x^{\alpha-1} dx \quad (\alpha > 0)$$

Propriedades:

- $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1) \quad (\alpha > 1)$
- $\Gamma(n) = (n-1)! \quad \alpha = n$
- $\Gamma(1) = 1$ e $\Gamma(1/2) = \sqrt{\pi}$

# Capítulo 5 – Distribuições contínuas

**Definição 5.13 – Distribuição gama**

Uma variável aleatória $X$ com função densidade dada por,

$$f(x|\alpha,\lambda) = \frac{\lambda^{\alpha} e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)}, x > 0\,, \alpha > 0, \lambda > 0$$

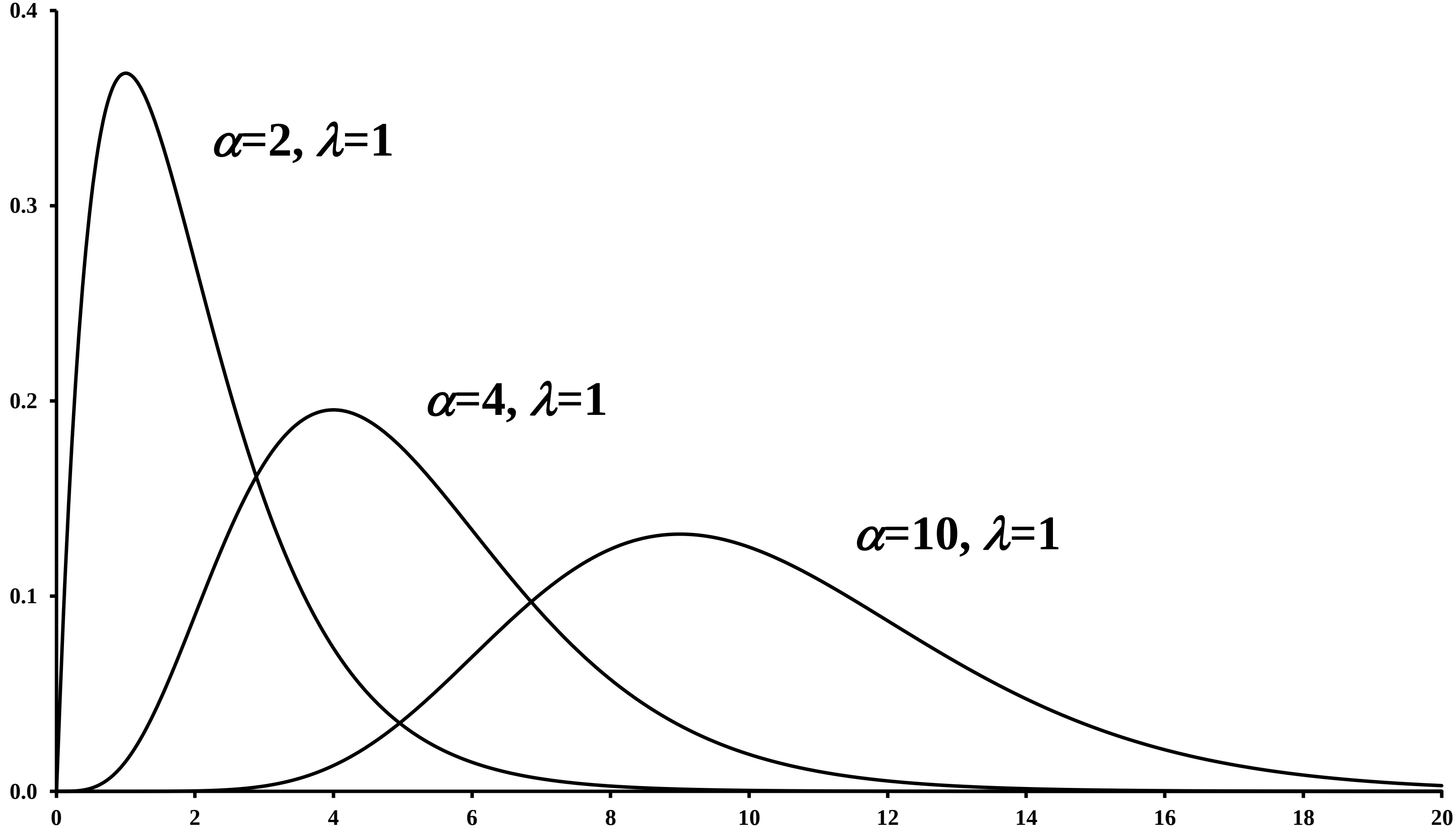
diz-se ter distribuição gama de parâmetros $\alpha$ e $\lambda$ .

Simbolicamente, $X \sim G(\alpha,\lambda)$

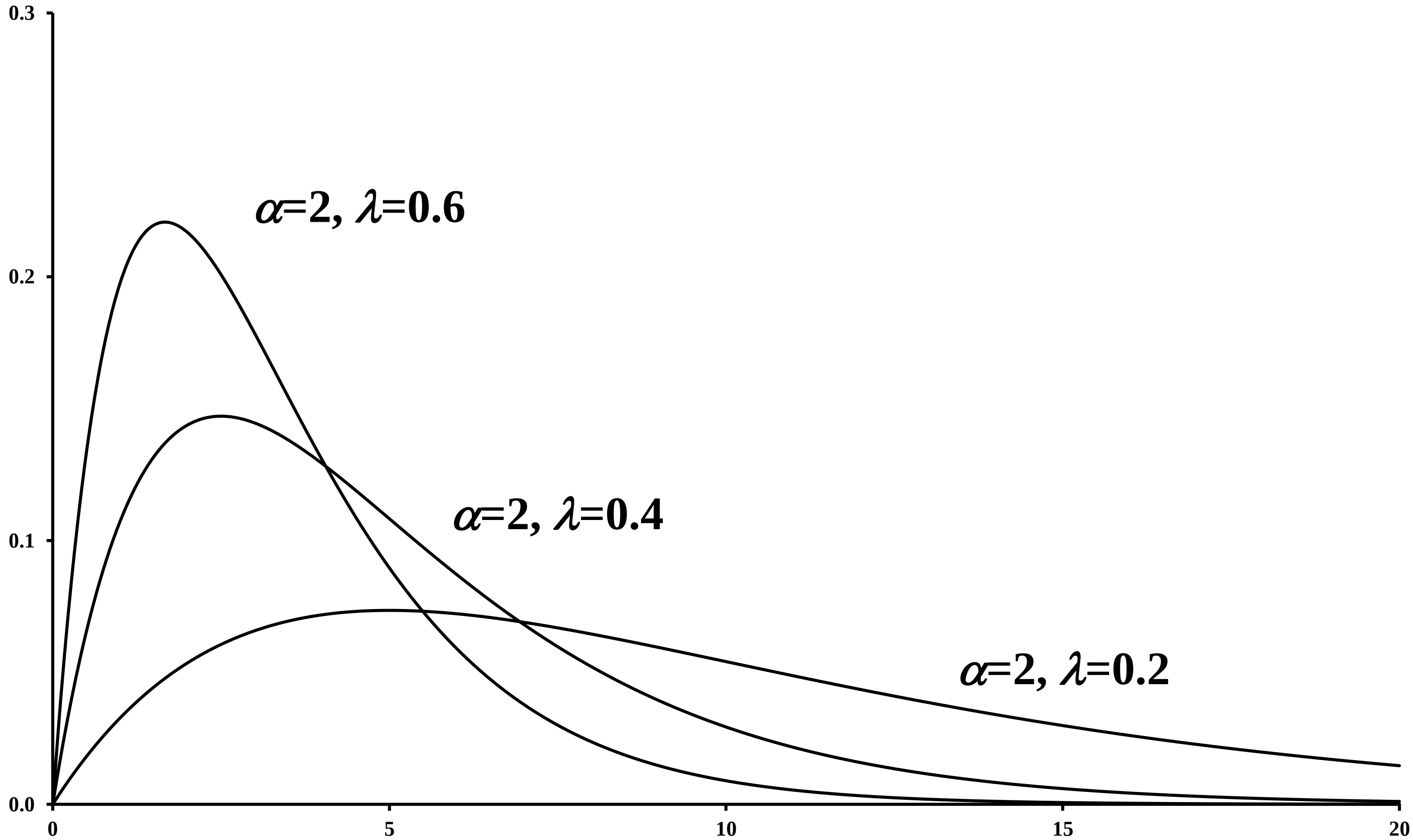
**Casos Particulares:**

- Quando o parâmetro $\alpha$ é inteiro, a dist. gama é conhecida por **dist. de Erlang**
- Quando ,$\alpha = 1$ temos a distribuição exponencial.
- Se, $\alpha = \frac{n}{2}, \lambda = \frac{1}{2}$ temos a distribuição do **Qui-quadrado**, $\chi^2_{(n)} = G\left(\frac{n}{2},\frac{1}{2}\right)$

# Capítulo 5 – Distribuições contínuas



# Capítulo 5 – Distribuições contínuas



# Capítulo 5 – Distribuições contínuas

$$E(X) = \frac{\alpha}{\lambda} \qquad Var(X) = \frac{\alpha}{\lambda^2} \qquad \gamma_1 = \frac{2}{\sqrt{\alpha}}$$

- **Mediana**: existe sempre, embora não se consiga explicitar analiticamente a sua expressão.

- Quando existem as três medidas de localização $(\alpha > 1)$, verifica-se a seguinte relação: $\mu_* < \mu_e < \mu$

- Não existe, em geral, solução analítica do integral correspon-dente à função de distribuição.

**Nota:** Não existem tabelas de valores da f. distribuição da Gama

# Capítulo 5 – Distribuições contínuas

**Teorema 5.5 –** Sejam $X_1, X_2$, duas variáveis aleatórias independentes. Então,

$$X_1 \sim G(\alpha_1, \lambda) \text{ e } X_2 \sim G(\alpha_2, \lambda) \Rightarrow X_1 + X_2 \sim G(\alpha_1 + \alpha_2, \lambda)$$

Generalizando para $k$ variáveis aleatórias independentes, tem-se,

$$X_i \sim G(\alpha_i, \lambda)(i = 1,2,\cdots,k) \Rightarrow \sum_{i=1}^{k} X_i \sim G\left(\sum_{i=1}^{k} \alpha_i, \lambda\right)$$

**Notas**: o parâmetro $\boldsymbol{\lambda}$ é o mesmo nas distribuições das parcelas.

A distribuição exponencial é um caso particular da distribuição gama ($\alpha = 1$);

A distribuição gama com $\alpha = n$ inteiro pode ser interpretada como a soma de $\boldsymbol{n}$ exponenciais independentes (teorema 5.5)

**Definição 5.14 – Distribuição do qui-quadrado**

Diz-se que a variável aleatória $X$ tem distribuição do qui-quadrado com $n$ graus de liberdade ($n$ inteiro positivo), quando a Respectiva função densidade é da forma,

$$f(x|n) = \frac{e^{-\frac{x}{2}} x^{\frac{n}{2}-1}}{2^{\frac{n}{2}}\Gamma(n/2)}, \quad x > 0, n > 0 \qquad (5.50)$$

Simbolicamente, $X \sim \chi^2_{(n)}$.

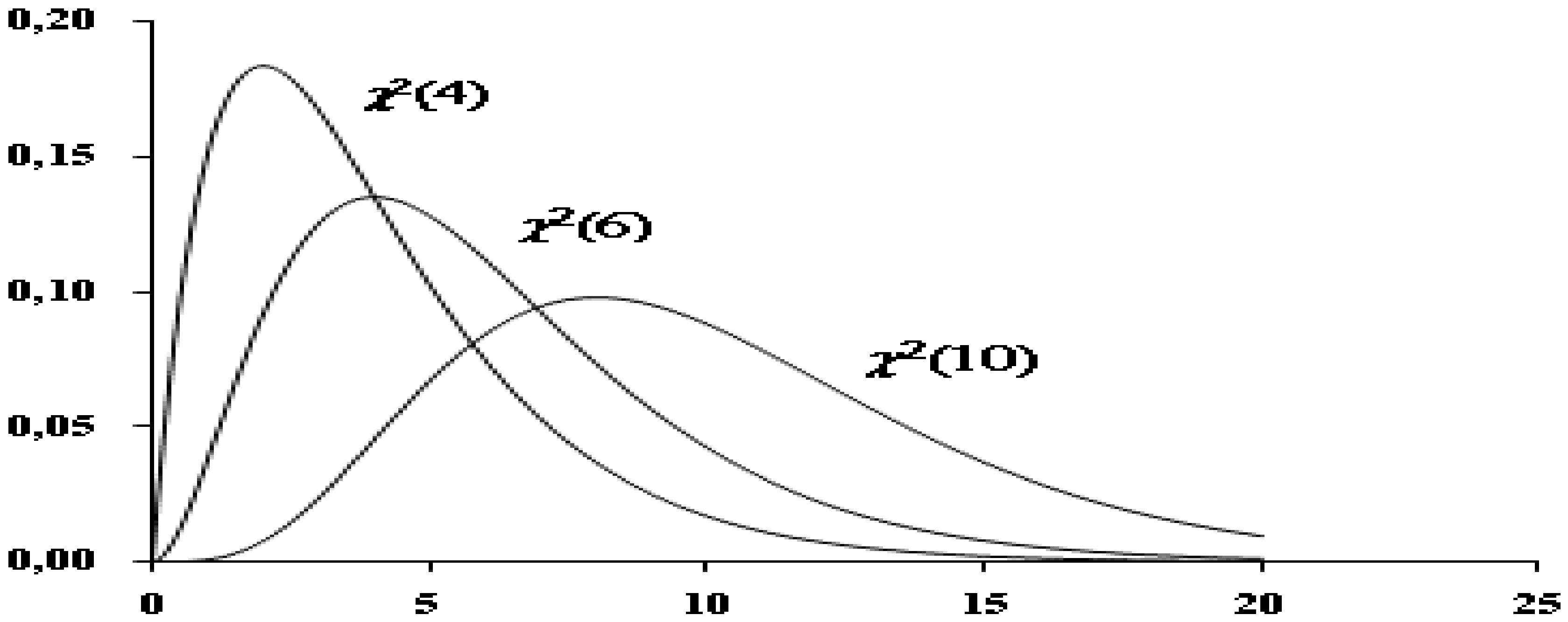
A distribuição do qui-quadrado é uma distribuição gama com

$\alpha = \frac{n}{2}$ e $\lambda = \frac{1}{2}$ ,i.é, $X \sim \chi^2_{(n)} \Leftrightarrow X \sim G\left(\frac{n}{2}, \frac{1}{2}\right)$

# Capítulo 5 – Distribuições contínuas

## Distribuição do qui-quadrado

$\boldsymbol{E(X) = n}$ ; $\boldsymbol{Var(X) = 2n}$ ; $\gamma_1 = \sqrt{\frac{8}{n}}$



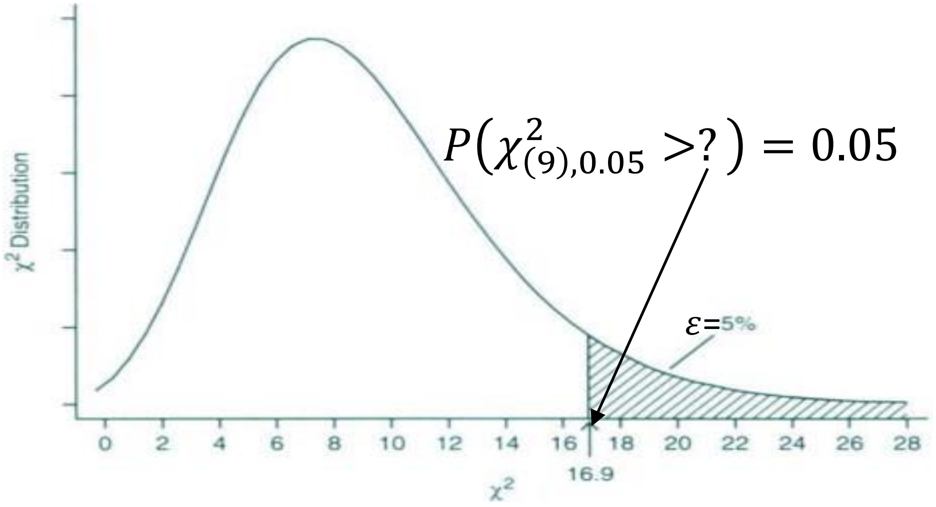
- Resultado importante

$$X \sim G(n, \lambda) \Leftrightarrow Y = 2\lambda X \sim \chi^2_{(2n)}$$

**Como calcular probabilidades?** Máquina de calcular ou computador

Tabelas

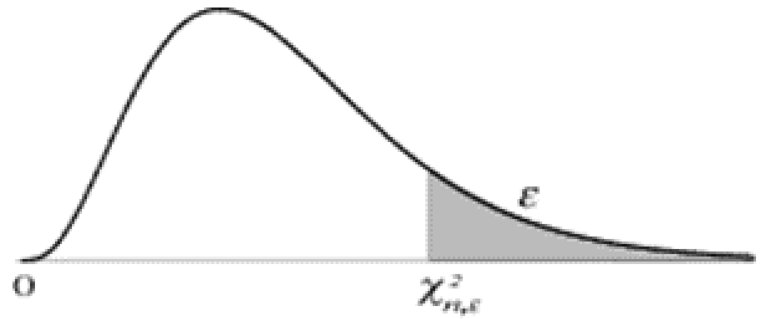
A tabela permite encontrar o valor $\chi^2_{(n),\varepsilon}$ tal que, $P\left(X > \chi^2_{(n),\varepsilon}\right) = \varepsilon$ para alguns valores de $\varepsilon$ e de $n$.



# Capítulo 5 – Distribuições contínuas

## TABELA 6 – DISTRIBUIÇÃO DO QUI-QUADRADO

$$\chi^2_{(n),\varepsilon}: P\left(X > \chi^2_{(n),\varepsilon}\right) = \varepsilon$$




| n \ ε | .995 | .990 | .975 | .950 | .900 | .750 | .500 | .250 | .100 | .050 | .025 | .010 | .005 | .001 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | .000 | .000 | .001 | .004 | .016 | .102 | .455 | 1.323 | 2.706 | 3.841 | 5.024 | 6.635 | 7.879 | 10.827 |
| **2** | .010 | .020 | .051 | .103 | .211 | .575 | 1.386 | 2.773 | 4.605 | 5.991 | 7.378 | 9.210 | 10.597 | 13.815 |
| **3** | .072 | .115 | .216 | .352 | .584 | 1.213 | 2.366 | 4.108 | 6.251 | 7.815 | 9.348 | 11.345 | 12.838 | 16.266 |
| **4** | .207 | .297 | .484 | .711 | 1.064 | 1.923 | 3.357 | 5.385 | 7.779 | 9.488 | 11.143 | 13.277 | 14.860 | 18.466 |
| **5** | .412 | .554 | .831 | 1.145 | 1.610 | 2.675 | 4.351 | 6.626 | 9.236 | 11.070 | 12.832 | 15.086 | 16.750 | 20.515 |
| **6** | .676 | .872 | 1.237 | 1.635 | 2.204 | 3.455 | 5.348 | 7.841 | 10.645 | 12.592 | 14.449 | 16.812 | 18.548 | 22.457 |
| **7** | .989 | 1.239 | 1.690 | 2.167 | 2.833 | 4.255 | 6.346 | 9.037 | 12.017 | 14.067 | 16.013 | 18.475 | 20.278 | 24.321 |
| **8** | 1.344 | 1.647 | 2.180 | 2.733 | 3.490 | 5.071 | 7.344 | 10.219 | 13.362 | 15.507 | 17.535 | 20.090 | 21.955 | 26.124 |
| **9** | 1.735 | 2.088 | 2.700 | 3.325 | 4.168 | 5.899 | 8.343 | 11.389 | 14.684 | 16.919 | 19.023 | 21.666 | 23.589 | 27.877 |
| **10** | 2.156 | 2.558 | 3.247 | 3.940 | 4.865 | 6.737 | 9.342 | 12.549 | 15.987 | 18.307 | 20.483 | 23.209 | 25.188 | 29.588 |

# Capítulo 5 – Distribuições contínuas

- Resultado importante: $\boldsymbol{X \sim G(n, \lambda) \Leftrightarrow Y = 2\lambda X \sim \chi^2_{(2n)}}$

**Exemplo 5.19 –** Seja $X$ – indemnizações pagas por certo risco por uma seguradora,

$$X \sim G(2; 0{,}008) \Leftrightarrow Y = 2\boldsymbol{\lambda} X = 2 * 0{,}008X = 0{,}016X \sim \chi^2_{(4)}$$

$$P(X > 500) = P(Y > 0{,}016 * 500) = P(Y > 8) \approx 0{,}1$$

Na tabela 6,$P(Y > 7.77944) = 0.1$ e $P(Y > 9.48773) = 0.05$

Valor em computador:$P(Y > 8) = chicdfrt(4, 8) = 0.0916$

- Para valores que não estão contemplados na tabela 6, pode utilizar-se para $n > 100$

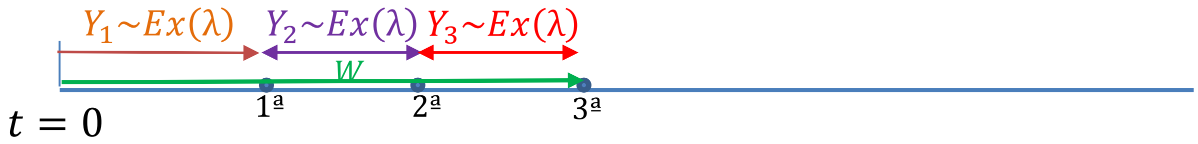
  $X \sim \chi^2_{(n)} \Rightarrow \sqrt{2X} - \sqrt{2n-1} \dot{\sim} N(0, 1)$

  onde o símbolo $\dot{\sim}$ significa distribuição aproximada ou assintótica.

# Capítulo 5 – Distribuições contínuas

3.b) Qual a probabilidade de se ter de esperar mais que 30 minutos até se observar a 3ª partícula?

$$W - tempo\ de\ espera\ até\ à\ observação\ da\ 3ª\ partícula$$




$$W = \sum_{i=1}^{3} Y_i \sim G(3, 0.1897) \Rightarrow 2\lambda^* W \sim \chi^2_{(2*3)}$$

$$P(W > 30) = P(2\lambda^* W > 2 * 0.1897 * 30) = P\left(\chi^2_{(6)} > 11.382\right) = 0.0773$$

c) Se já passaram 15 minutos desde que se observou uma partícula qual a probabilidade de se ter de esperar ainda pelo menos 5 minutos pela observação da próxima?

$$P(Y > 20 | Y > 15) = P(Y > 5)$$

# Capítulo 5 – Distribuições contínuas

- **Teorema 5.6 –** Sejam $X_1, X_2$, duas variáveis aleatórias independentes. Então, se

$$X_1 \sim \chi^2_{(n_1)} \text{ e } X_2 \sim \chi^2_{(n_2)} \Rightarrow X_1 + X_2 \sim \chi^2_{(n_1+n_2)}$$

Generalizando para $k$ variáveis aleatórias independentes, tem-se:

$$X_i \sim \chi^2_{(n_i)} (i = 1,2,\cdots,k) \Rightarrow Y = \sum_{i=1}^{k} X_i \sim \chi^2_{\left(n=\sum_{i=1}^{k} n_i\right)}$$

**Teorema 5.7 –** Seja $X$ uma variável aleatória. Então, se

$$\boldsymbol{X \sim N(0,1) \Rightarrow Y = X^2 \sim \chi^2_{(1)}}$$

***Dem*.**: livro

# Capítulo 5 – Distribuições contínuas

**Teorema 5.8 –** Se as variáveis aleatórias $X_i \quad i = 1, 2, \cdots, n$, são Independentes,

$$\boldsymbol{X_i \sim N(0, 1) \Rightarrow \sum_{i=1}^{n} X_i^2 \sim \chi^2_{(n)}}$$

O **teorema 5.8** pode também ser enunciado da seguinte maneira:
**Se as variáveis aleatórias** $X_i \quad i = 1, 2, \cdots, n$**, são independentes e** ***normais*****...**

$$\boldsymbol{X_i \sim N\left(\mu_i, \sigma_i^2\right) \Rightarrow \sum_{i=1}^{n} \left(\frac{X_i - \mu_i}{\sigma_i}\right)^2 \sim \chi^2_{(n)}}$$

# Capítulo 5 – Distribuições contínuas

## 5.5 - Teorema do limite central

O Vice-presidente para a área financeira de uma grande empresa de seguros de saúde  deseja monitorizar os pagamentos diários de participações de seguros para determinar se o valor médio em euros dessas participações é estável, aumenta ou decresce. O valor das participações individuais varia muito, de um dia para o outro e seria ingénuo tirar conclusões com base nas variações diárias. Existem, no entanto, momentos em que as mudanças tornam-se substanciais e devem ser objecto de intervenção. Registos históricos mostram que as participações de seguros de saúde são altamente enviesadas e que sua distribuição possui uma média de 60 e um desvio padrão de 225 euros. Usando essa informação, é lhe pedido que  desenvolva um sistema de monitorização das participações baseado numa amostra aleatória de 100 participações. A empresa estabeleceu um intervalo de aceitação de 95% para variação do valor das participações.

# Capítulo 5 – Distribuições contínuas

## 5.5 - Teorema do limite central

- Um dos resultados mais importantes da teoria da probabilidade.

- Introduz a noção de **distribuição aproximada** ou **assimptótica**.

# Capítulo 5 – Distribuições contínuas

## 5.5 - Teorema do limite central

**Teorema 5.10 Teorema do limite central (Lindberg-Levy)**

Dada a sucessão de variáveis aleatórias $iid, X_1, X_2, \cdots, X_n, \cdots$, com média $\mu$ e variância $\sigma^2$(finita), então, quando $n \to +\infty$, a função de distribuição da variável aleatória,

$$Z_n = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \dot{\sim} N(0,1)$$

Diz-se que $Z_n$ tende para uma função de distribuição $N(0,1)$ , ou seja, a distribuição assintótica de $Z_n$ é a $N(0,1)$.

# Capítulo 5 – Distribuições contínuas

## Interpretação prática do TLC

A conclusão do teorema anterior, desde que $n$ grande, pode exprimir-se na forma alternativa,

$$\lim_{n\to+\infty} P(Z_n \leq z) = \lim_{n\to+\infty} P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \leq z\right) = P(Z \leq z) = \boldsymbol{\Phi}(z)$$

Ou ainda

$$P(Z_n \leq z) = P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \leq z\right) \approx \boldsymbol{\Phi}(z)$$

## Notas:

**1.** Distribuição de $X_i$ pode ser **discreta** ou **contínua**.

**2. O T.L.C.** pode aplicar-se desde que **se conheçam a média e variância** da v.a. $X_i$, **mesmo que a sua distribuição seja desconhecida**

# Capítulo 5 – Distribuições contínuas

O que é *n* grande?  Depende da distribuição de $X_i$.

Distribuições simétricas e unimodais tornam a convergência mais rápida e melhor a aproximação.

Aconselhável $n \geq 30$, salvo nalguns casos (convergência rápida).

Alguns cuidados **quando a distribuição de $X_i$ é discreta**: **Correcção de continuidade**.

# Capítulo 5 – Distribuições contínuas

O Vice-presidente para a área financeira de uma grande empresa de seguros de saúde deseja monitorizar os pagamentos diários de participações de seguros para determinar se o valor médio em euros dessas participações é estável, aumenta ou decresce. O valor das participações individuais varia muito, de um dia para o outro e seria ingénuo tirar conclusões com base nas variações diárias. Existem, no entanto, momentos em que as mudanças tornam-se substanciais e devem ser objecto de intervenção. Registos históricos mostram que as participações de seguros de saúde são altamente enviesadas e que sua distribuição possui uma média de 60 e um desvio padrão de 225 euros. Usando essa informação, é lhe pedido que desenvolva um sistema de monitorização das participações baseado numa amostra aleatória de 100 participações. A empresa estabeleceu um intervalo de aceitação de 95% para variação do valor das participações.

# Capítulo 5 – Distribuições contínuas

$X_i$ – valor, em euros, de uma participação individual

$X_i: E(X) = 60; Var(X) = 225; \sigma_X = 15; X_i \sim F(X) = ?$

Amostra de $n = 100$ participações

$$\sum_{i=1}^{100} X_i - valor\ total\ das\ participações\ na\ amostra$$

$$a, b = ?: P\left(a < \sum_{i=1}^{n} X_i < b\right) = 0.95$$

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

$$P\left(\frac{\boldsymbol{-a - 100 * 60}}{\boldsymbol{15\sqrt{100}}} < \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} < \frac{\boldsymbol{a - 100 * 60}}{\boldsymbol{15\sqrt{100}}}\right) = 0.95$$

$$P\left(\underbrace{\frac{\boldsymbol{-a - 100 * 60}}{\boldsymbol{15\sqrt{100}}}}_{-z\varepsilon/_2} < Z < \underbrace{\frac{\boldsymbol{a - 100 * 60}}{\boldsymbol{15\sqrt{100}}}}_{z\varepsilon/_2}\right) = 0.95$$

$$z\varepsilon/_2 = P\left(Z > z\varepsilon/_2\right) = 0.025$$

$$\Rightarrow z\varepsilon/_2 = 1.96 = \frac{\boldsymbol{a - 100 * 60}}{\boldsymbol{15\sqrt{100}}} \Leftrightarrow \boldsymbol{a = 6002,94} \qquad -1.96 = \frac{\boldsymbol{a - 100 * 60}}{\boldsymbol{15\sqrt{100}}} \Leftrightarrow \boldsymbol{a = 5997,6}$$

$$P\left(a < \sum_{i=1}^{n} X_i < b\right) = 0.95 \qquad a, b:\ (5997.6, 6002.4)$$

# Capítulo 5 – Distribuições contínuas

$X_i$ – valor, em euros, de uma participação individual

$X_i: E(X) = 60; Var(X) = 225; \sigma_X = 15$ Amostra de $n = 100$ participações

$$\sum_{i=1}^{100} X_i - valor\ total\ das\ participações\ na\ amostra$$

$$a, b =?: P\left(a < \sum_{i=1}^{n} X_i < b\right) = 0.95 \qquad \sum_{i=1}^{n} X_i \dot{\sim} N(100 * 60, 100 * 225)$$

$$P\left(-\boldsymbol{a} < \sum_{i=1}^{n} X_i < \boldsymbol{a}\right) = 0.95 \quad \Rightarrow a = invnorm(0.975, 6000, 150)$$

$$P\left(a < \sum_{i=1}^{n} X_i < b\right) = 0.95 \qquad a, b:\ (5997.6, 6002.4)$$

**Exemplo 5.22 –** A procura diária a satisfazer (unidade: $100\ Kg$) é uma v.a. $X$ com média $40$ e variância $25$. A produção anual planeada é de $11\ 500$, calcular a probabilidade de haver procura anual excedentária (ano: $289$ dias úteis). $X_i$ -procura no dia $i$

$$E(X_i) = \mu = 40 \text{ e } Var(X_i) = \sigma^2 = 25; X_i \sim F(X) = ?$$

$$\sum_{i=1}^{n} X_i \dot{\sim} N(289 * 40, 289 * 225)$$

$$P\left(\sum_{i=1}^{289} X_i > 11500\right) = P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} > \frac{11500 - 289 * 40}{5 * \sqrt{289}}\right) = \mathrm{P}(Z > 0.71)$$

Tabela $\approx 1 - P(Z_n \leq -0{,}71) = 0{,}7611$

Máquina calcular $\approx 1 - normcdf\left(-0.71, 289 * 40, 5 * \sqrt{289}\right) = 0{,}7611$

$$\approx 1 - normcdf\left(11500, 99999, 289 * 40, 5 * \sqrt{289}\right) = 0{,}7611$$

# Capítulo 5 – Distribuições contínuas

**Exemplo 5.23 –** Do exemplo anterior pretende determinar-se a produção $Q$ que deve ser planeada de modo a cobrir a procura anual com probabilidade 0.99.

$$Q =?: P\left(\sum_{i=1}^{289} X_i < Q\right) = 0.99 \Leftrightarrow P\left(\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} < \frac{Q - 289 * 40}{5 * \sqrt{289}}\right) = 0{,}99$$

Tabela $\approx \boldsymbol{\Phi}\left(\frac{Q - 289 * 40}{5 * \sqrt{289}}\right) = 0{,}99 \Leftrightarrow \frac{Q - 289 * 40}{5 * \sqrt{289}} = 2{,}327$

$$\Leftrightarrow Q = 11758(10^2 kgs)$$

Máquina de calcular

$$Q =?: P\left(\sum_{i=1}^{289} X_i < Q\right) = 0.99 \Leftrightarrow Q = invnorm(0.99{,}289 * 40, 5 * \sqrt{289})$$

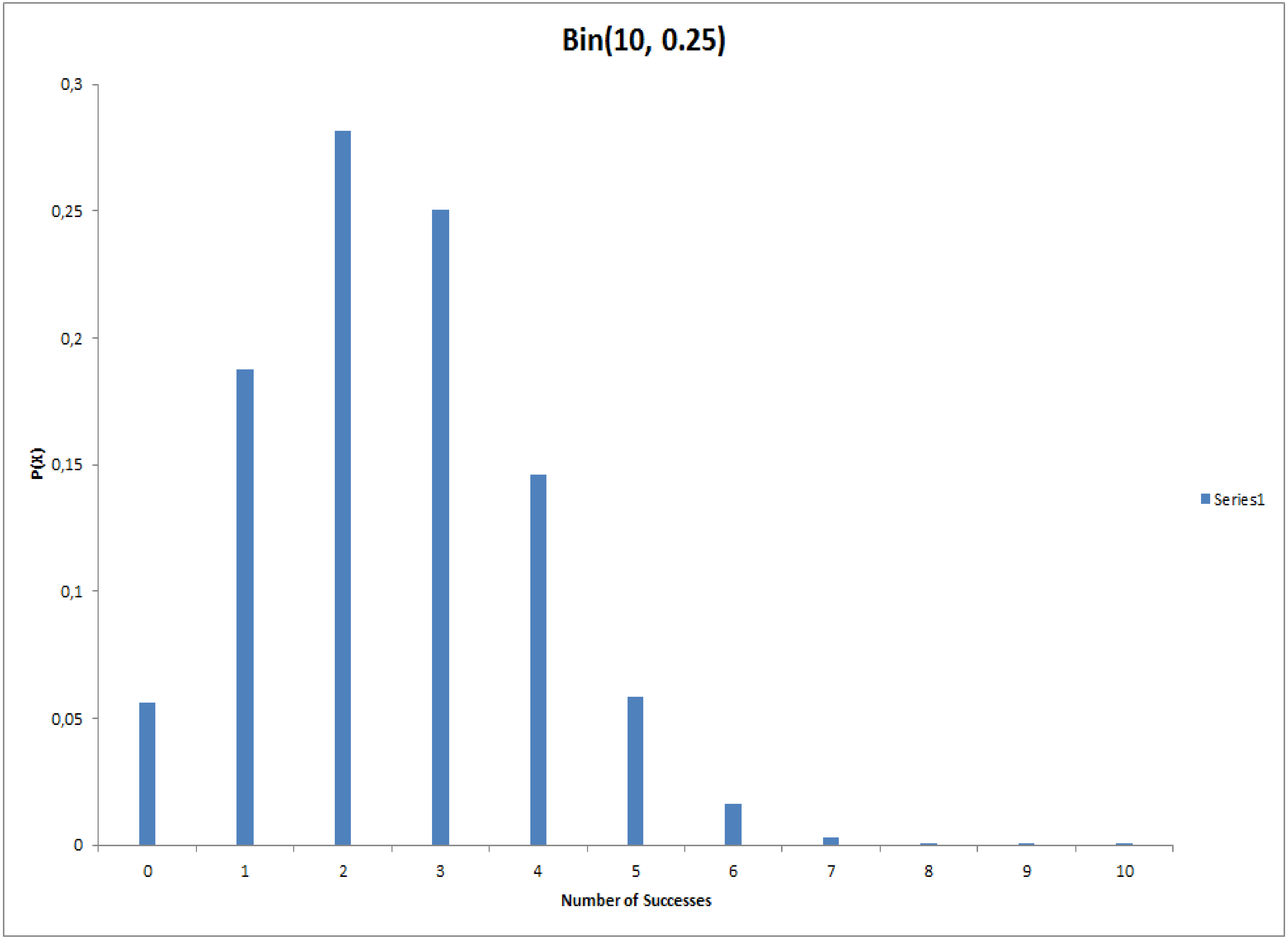
# T. Limite Central: Aplicação a distribuições discretas

**Corolário 5.3 (T. de De Moivre-Laplace)** – Dada a sucessão de variáveis aleatórias $iid$, $X_1, X_2, \cdots, X_n, \cdots$, com distribuição de Bernoulli de média $E(X_i) = \mu = \theta$ e $Var(X_i) = \sigma^2 = \theta(1-\theta)$, tem-se,

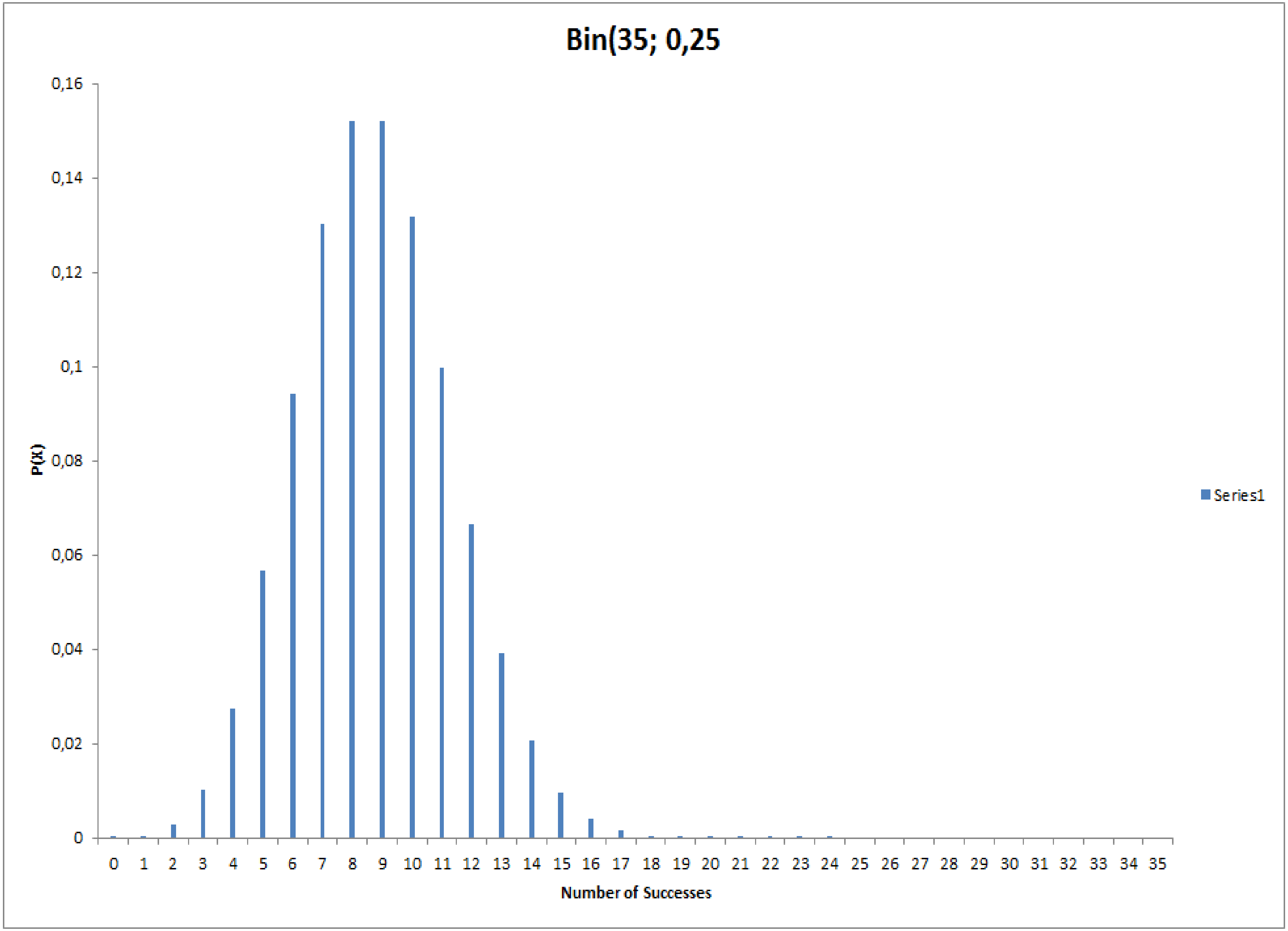
$$Z_n = \frac{\sum_{i=1}^{n} X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \dot{\sim} N(0,1)$$

**Nota:** $\sum_{i=1}^{n} X_i \sim \boldsymbol{Bi(n,\theta)}$. Quando $n$ é grande, utilizar o corolário.

Mas... **com Correcção de Continuidade.**



Bin(10, 0.25)
0,3
0,25
0,2
0,15
0,1
0,05
0
P(X)
0
1
2
3
4
5
6
7
8
9
10
Number of Successes
Series1


Bin(35; 0,25
0,16
0,14
0,12
0,1
0,08
0,06
0,04
0,02
0
P(X)
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35
Number of Successes
Series1

# Capítulo 5 – Distribuições contínuas

$\boldsymbol{Y = \sum_{i=1}^{n} X_i \sim Bi(n, \theta)}$ com $n$ grande

$P(a \leq Y \leq b) = ? \quad a, b\ inteiros, \quad 0 \leq a < b \leq n$

- Resposta exacta: $$P(a \leq Y \leq b) = \sum_{y=a}^{b} \binom{n}{y} \theta^y (1-\theta)^{n-y}$$

- Resposta aproximada sem correcção de continuidade: $$\frac{\sum_{i=1}^{n} X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \dot{\sim} N(0,1)$$

$$P(a \leq Y \leq b) \approx \mathbf{\Phi}\left(\frac{b - n\theta}{\sqrt{n\theta(1-\theta)}}\right) - \mathbf{\Phi}\left(\frac{a - n\theta}{\sqrt{n\theta(1-\theta)}}\right)$$

# Capítulo 5 – Distribuições contínuas

## *Se $X$ é uma variável discreta tem-se:*

$$P(a \leq X \leq b) = P(a - d_1 \leq X \leq b + d_2),\ d_i \subset \mathbb{R} : 0 \leq d_i < 1\ (i = 1{,}2)$$

Os casos extremos são:

$$d_1 = 1, d_2 = 0\ \Rightarrow P(a - 1 \leq X \leq b)$$

$$d_1 = 0, d_2 = 1\ \Rightarrow P(a \leq X \leq b + 1)$$

$$d_1 = 1, d_2 = 1\ \Rightarrow P(a - 1 \leq X \leq b + 1)$$

e estas probabilidades são todas iguais porque $X$ *é uma variável discreta*

Ao calcular esta probabilidade usando o TLC, obtêm-se diferentes valores para diferentes valores de $d_i$ porque se utiliza uma distribuição contínua. Teríamos assim, várias aproximações. Como escolher a melhor?

Embora não exista solução óptima, é prática comum fazer $d_1 = d_2 = \frac{1}{2}$ conhecida como **correcção de continuidade**.

# • Correcção de continuidade

$$P(a \leq Y \leq b) = P\left(a - \frac{1}{2} \leq Y \leq b + \frac{1}{2}\right) \approx \mathbf{\Phi}\left(\frac{b + \frac{1}{2} - n\theta}{\sqrt{n\theta(1-\theta)}}\right) - \mathbf{\Phi}\left(\frac{a - \frac{1}{2} - n\theta}{\sqrt{n\theta(1-\theta)}}\right)$$

A **correcção de continuidade** **aplica-se** **sempre** **quando se** **aproxima uma distribuição discreta por uma distribuição contínua**

# Capítulo 5 – Distribuições contínuas

- Cálculo de probabilidades com a binomial (regra prática):

  Sempre que possível utilizar directamente a distribuição binomial.

- Quando necessário, o cálculo aproximado das probabilidades deve atender aos seguintes casos:

  - $Se\ 0{,}1 < \theta < 0{,}9 \rightarrow (ou\ n\theta \geq 5\ )$ aproximar usando o TLC, <u>utilizando a correcção de continuidade</u>.

  - Se $\theta \leq 0{,}1$ ( $n$ muito grande)→ aproximar usando a Lei Acontecimentos Raros.

  - Se $\theta \geq 0{,}9$ → aproximar usando a Lei Acontecimentos Raros, considerando o respectivo acontecimento complementar.

# Capítulo 5 – Distribuições contínuas

A gestão do sistema da rede de Pestana Hotéis acompanha a proporção de reservas feitas através da Internet para ajustar a sua publicidade e sistema de reservas web. Dados históricos mostram que a proporção de reservas pela Internet, por mês, é de 0,2.

a) Tendo sido seleccionada uma amostra de 64 clientes, qual a probabilidade de mais de 30 terem feito a reserva pela Internet?

X – reserve feita pela internet $\sim B(1, 0.2)$ $\quad \sum_{i=1}^{64} X_i \dot{\sim} N(64 * 0.2, 64 * 0.2 * (1-0.2))$

$$P\left(\sum_{i=1}^{64} X_i > 30\right) = 1 - P\left(\sum_{i=1}^{64} X_i \leq 30\right) = 1 - P\left(\sum_{i=1}^{64} X_i \leq 30 + 0.5\right) = 1 - P\left(\sum_{i=1}^{64} X_i \leq 30,5\right)$$

$$= 1 - normcdf(-99999, 30.5, 12.8, 3.2) = 1 - 1 = 0$$

$$P\left(\sum_{i=1}^{64} X_i \leq 30,5\right) = P\left(\frac{\sum_{i=1}^{64} X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \leq \frac{30.5 - 12.8}{\sqrt{64 * 0.2 * 0.8}}\right) = \mathbf{\Phi}(5.53) \approx 1$$

b) Qual o número máximo de clientes que fará reservas pela Internet com uma probabilidade de 90%?

$$n = 64: P\left(\sum_{i=1}^{64} X_i \leq a\right) = 0.9$$

$$a = invnorm(0.9, 12.8, 3.2) = 16.9 + 0.5 = 17.3 \quad R: a = 17$$

$$P\left(\frac{\sum_{i=1}^{64} X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \leq \frac{a + 0.5 - 12.8}{\sqrt{64 * 0.2 * 0.8}}\right) = \mathbf{\Phi}^{-1}(0.9)$$

$z_\varepsilon = P(Z > z_\varepsilon) = 0.1 \Rightarrow z_\varepsilon = 1.282$

# Capítulo 5 – Distribuições contínuas

**Corolário 5.4 –** Se $X$ é uma variável com distribuição de Poisson, $X \sim Po(\lambda)$ , então, $\lim_{n \to +\infty} P\left(\frac{X-\lambda}{\sqrt{\lambda}} \leq x\right) = \phi(x)$

Demonstração: livro

# Capítulo 5 – Distribuições contínuas

## 5.6 - Distribuição Normal bidimensional

**Definição 5.21 – Distribuição normal bidimensional**

Se a variável aleatória bidimensional, (X,Y), tem função densidade da forma,

$$f\left(x, y | \mu_X, \mu_y, \sigma_X^2, \sigma_Y^2, \rho\right) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \times$$
$$\times\, exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_X}{\sigma_X}\right)^2 - \right.\right.$$

# Capítulo 5 – Distribuições contínuas

Pode demonstrar-se [Murteira (1990a)] que,

- $\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)dxdy = 1$
- $E(X) = \mu_x, E(Y) = \mu_y$
- $Var(X) = \sigma_X^2, Var(Y) = \sigma_Y^2$
- $Cov(X,Y) = \rho\sigma_X\sigma_Y$ ($\rho$ é o coeficiente de correlação)

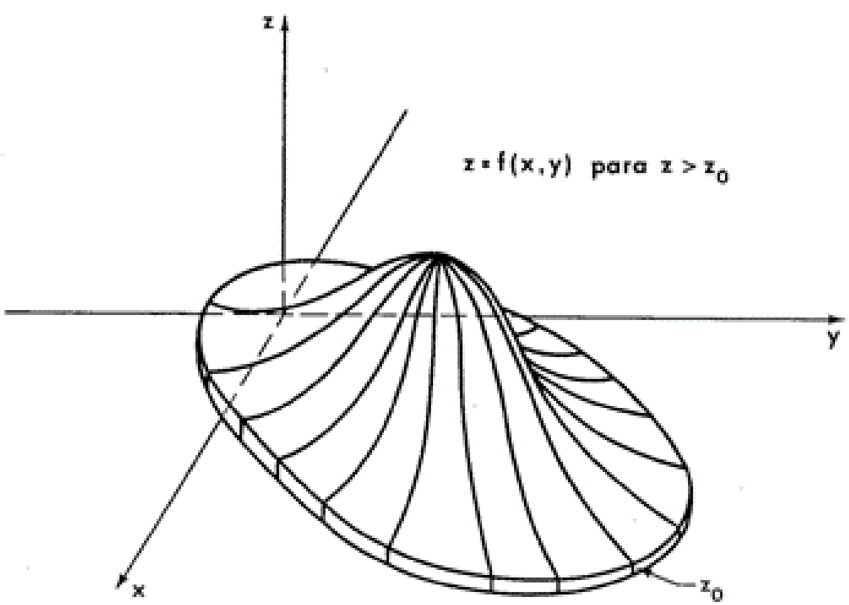
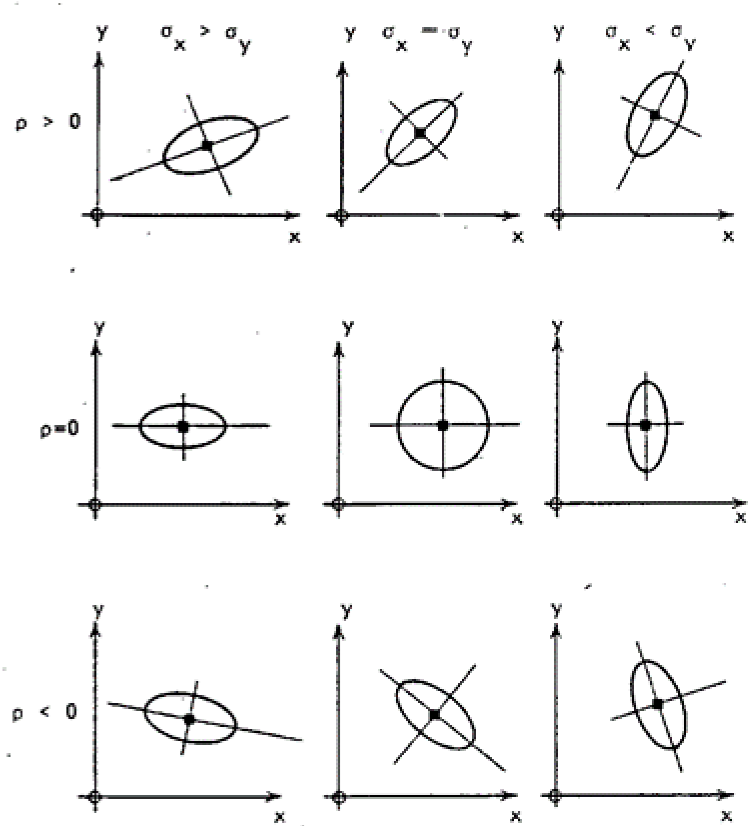


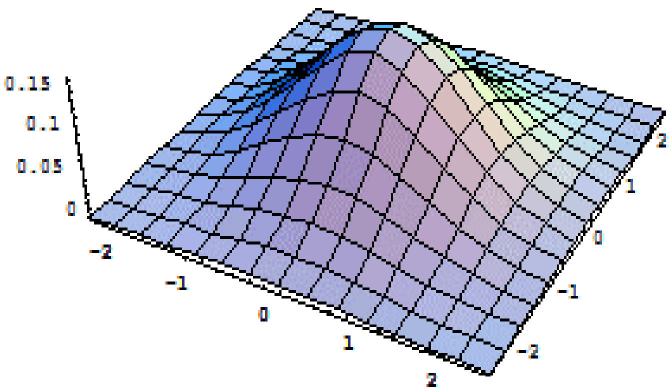

**Fig. 5.10 –** Função densidade da distribuição normal bidimensional

# Capítulo 5 – Distribuições contínuas

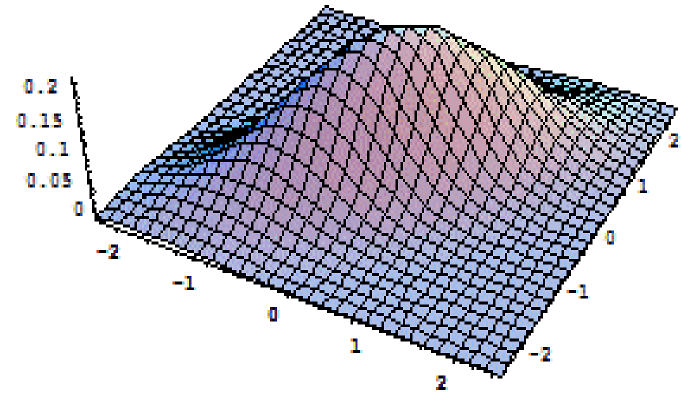
- A figura pode ser completada estudando as curvas de nível da função densidade, isto é, as intersecções da superfície $z = f(x, y)$ com planos paralelos ao plano $x0y$ e respectivas projecções neste plano (elipses de contorno).



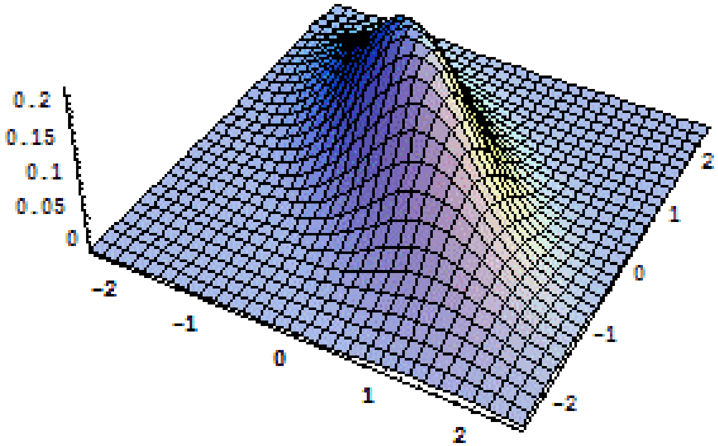
# Capítulo 5 – Distribuições contínuas




$\rho = 0$ **(Covariância nula)**




$\rho > 0$ **(Covariância positiva)**




$\rho < 0$ **(Covariância negativa)**

# Capítulo 5 – Distribuições contínuas

Caso particular: **Distribuição normal bidimensional estandardizada** $(\mu_X = \mu_Y = 0;\ \sigma_X^2 = \sigma_Y^2 = 1)$

$$f(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} exp\left\{-\frac{1}{2(1-\rho^2)}(x^2 - 2\rho xy + y^2)\right\}$$

Funções densidade marginais:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} exp\left\{-\frac{1}{2\sigma_X^2}(x-\mu_X)^2\right\}$$

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma_Y^2}} exp\left\{-\frac{1}{2\sigma_Y^2}(y-\mu_Y)^2\right\}$$

# Capítulo 5 – Distribuições contínuas

**Teorema 5.11** – Se  é uma variável aleatória normal bidimensional então $X$ e $Y$ são independentes se e só se $X$ e $Y$ forem não correlacionadas $(\rho = 0)$.

Dem: Ver livro

Funções densidade condicionadas (continuam a ser distribuições normais):

$$X \mid Y = y \sim N(\mu_X + \rho \frac{\sigma_X}{\sigma_Y}(y - \mu_Y);\ \sigma_X^2(1-\rho^2))$$

$$f_{X|Y=y}(x) = \frac{1}{\sqrt{2\pi\sigma_X^2(1-\rho^2)}} \times \exp\left\{ -\frac{1}{2\sigma_X^2(1-\rho^2)} \left[ x - \left( \mu_X + \rho \frac{\sigma_X}{\sigma_Y}(y - \mu_Y) \right) \right]^2 \right\}$$

$$Y \mid X = x \sim N(\mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X);\ \sigma_Y^2(1-\rho^2))$$

$$f_{Y|X=x}(y) = \frac{1}{\sqrt{2\pi\sigma_Y^2(1-\rho^2)}} \times \exp\left\{ -\frac{1}{2\sigma_Y^2(1-\rho^2)} \left[ y - \left( \mu_Y + \rho \frac{\sigma_Y}{\sigma_X}(x - \mu_X) \right) \right]^2 \right\}$$

# Capítulo 5 – Distribuições contínuas

## Valores esperados e variâncias condicionados

$$\mu_X = E(Y|X = x) = \mu_Y + \rho\,\frac{\sigma_Y}{\sigma_X}(x - \mu_X);$$

$$\mu_Y = E(X|Y = y) = \mu_X + \rho\,\frac{\sigma_X}{\sigma_Y}\left(y - \mu_y\right);$$

$$\sigma_X^2(x) = Var(Y|X = x) = \sigma_Y^2(1 - \rho^2);$$

$$\sigma_Y^2(y) = Var(X|Y = y) = \sigma_X^2(1 - \rho^2)$$

- Observações:
  - O valor esperado condicionado de $X$ é uma função linear de *y*.
  - A variância de $X$ condicionada por $Y = y$ não depende de *y* e é menor do que a variância de *X* (sem ser condicionada). Esta redução será tanto maior quanto mais correlacionadas forem as variáveis *X* e *Y*.