# Estatística I

## Licenciatura em Economia
## 2.º Ano/2.º Semestre
## 2023/2024

# Aulas Teóricas N.$^{os}$ 19 e 20 (Semana 11)

**Docente**: Elisabete Fernandes
**E-mail**: efernandes@iseg.ulisboa.pt

https://doity.com.br/estatistica-aplicada-a-nutricao

https://basiccode.com.br/produto/informatica-basica/

# Conteúdos Programáticos

| Aulas Teóricas (Semanas 1 a 3) | Aulas Teóricas (Semanas 3 a 5) | Aulas Teóricas (Semanas 5 a 7) | Aulas Teóricas (Semanas 8 a 13) |
| --- | --- | --- | --- |
| • **Capítulo 2**: Probabilidades | • **Capítulo 3**: Variáveis Aleatórias Unidimensionais | • **Capítulo 4**: Variáveis Aleatórias Multidimensionais | • **Capítulo 5**: Distribuições Teóricas<br>• **Capítulo 6**: Amostragem. Distribuições por Amostragem. |

**Material didático:** Exercícios do Livro Murteira et al (2015), Formulário e Tabelas Estatísticas

**Bibliografia**: B. Murteira, C. Silva Ribeiro, J. Andrade e Silva, C. Pimenta e F. Pimenta; *Introdução à Estatística*, 2ª ed., Escolar Editora, 2015.

https://cas.iseg.ulisboa.pt

# 1

# Distribuição t-Student

Variáveis Aleatórias Contínuas

# Distribuição t-Student

- A distribuição t-Student é uma distribuição de probabilidade contínua.

- A distribuição t-Student geralmente surge quando temos uma população com variância desconhecida (e tem de ser estimada a partir dos dados recolhidos) e uma amostra de dimensão pequena ($n < 30$).

- A distribuição t-Student é dada pelo quociente entre uma normal reduzida e a raiz quadrada de uma qui-quadrado dividida pelo respectivo número de graus de liberdade.

*I.e.*, se $Z \sim N(0; 1)$ e $Y \sim \chi^2_{(n)}$, duas variáveis aleatórias independentes, então:

$$\boxed{T = \frac{Z}{\sqrt{\frac{Y}{n}}} \cap t_{(n)}}$$

- Se X tem distribuição t-Student com n graus de liberdade, escreve-se:

$$\boxed{X \sim t_{(n)}}$$

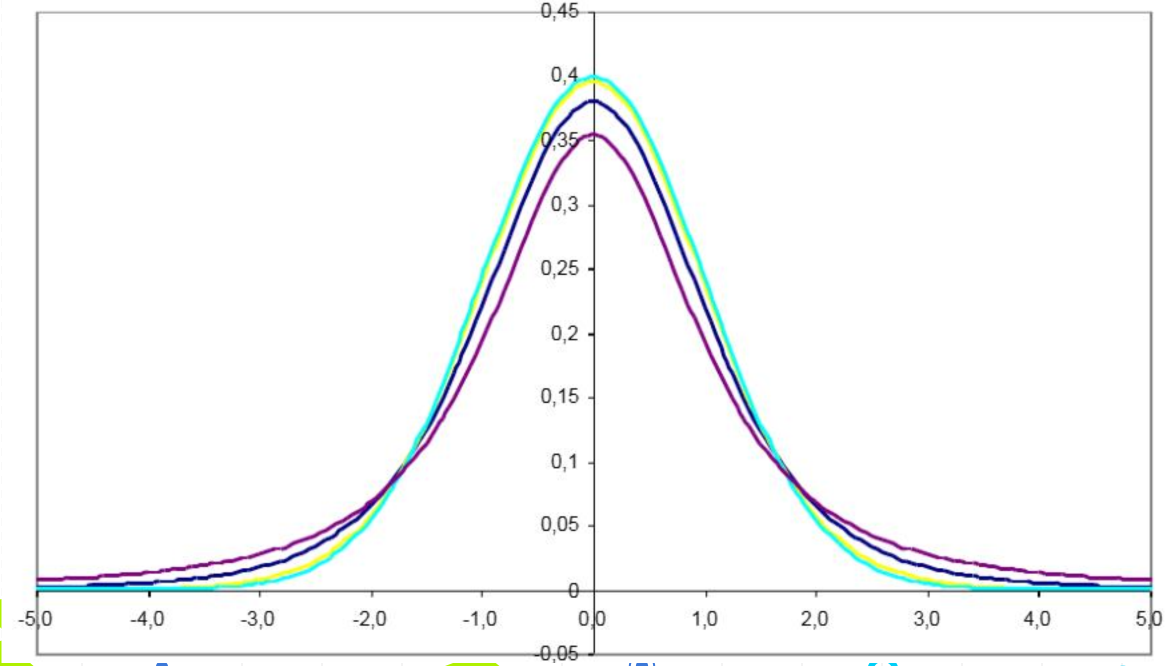
# Distribuição t-Student

- A distribuição t-Student tem um parâmetro: n – o nº de graus de liberdade.
- É uma distribuição simétrica.

- $E[X] = 0$
- $Var[X] = n/(n-2)$, se $n > 2$

Distribuição Qui-quadrado, Distribuição t-Student e Distribuição F-Snedecor - Distribuições - Studocu

# Distribuição t-Student

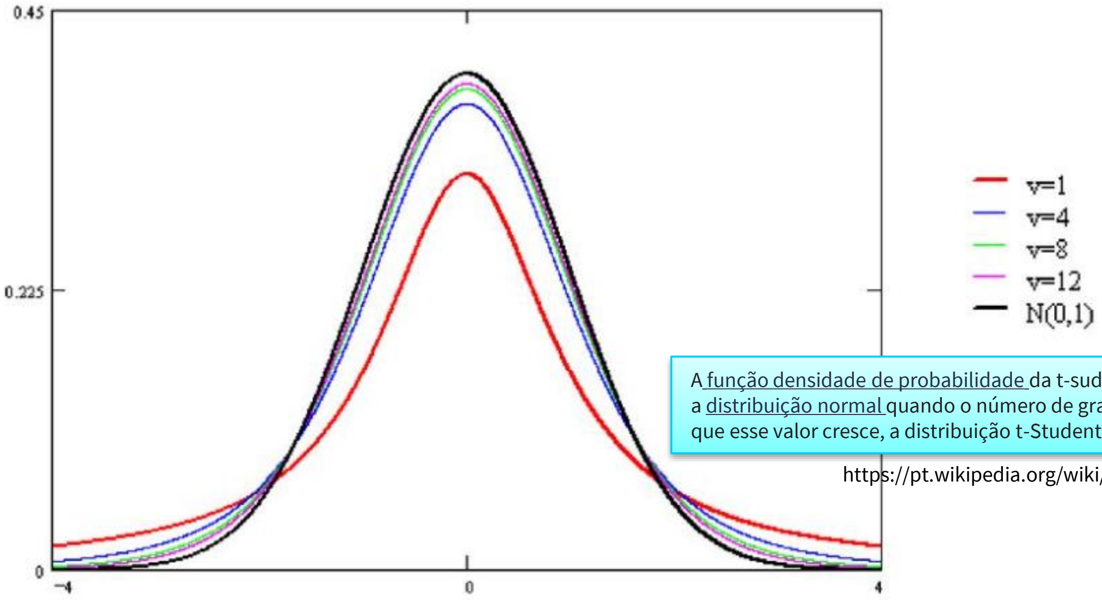



[Distribuição Qui-quadrado, Distribuição t-Student e Distribuição F-Snedecor - Distribuições - Studocu]

## T-Student

- Se a **variável tem distribuição Normal na população**, ou a amostra é suficientemente grande, mas não conhecemos o desvio da população, só da amostra, então ...
- ... A média amostral se distribui conforme uma **t-Student**
- **...** A distribuição t-Student depende dos graus de liberdade (n-1), que denotamos por ν

$$\frac{\overline{X}-\mu_0}{\frac{s}{\sqrt{n}}} \cap t_{(n-1)}$$




A função densidade de probabilidade da t-sudent detém caudas mais pesadas que a distribuição normal quando o número de graus de liberdade é pequeno e à medida que esse valor cresce, a distribuição t-Student aproxima-se da normal.

https://pt.wikipedia.org/wiki/Distribui%C3%A7%C3%A3o_t_de_Student

https://slideplayer.com.br/slide/14561961

# Distribuição t-Student

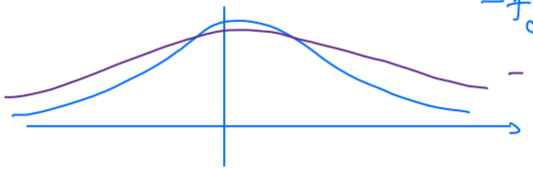
- A distribuição t-Student varia de acordo com os graus de liberdade. Isto significa que a sua curva depende da dimensão da amostra, n.
- Está tabelada para algumas probabilidade e $n \leq 30$, $n = 40$, $n = 60$, $n = 120$ e $n = \infty$.
- Quando $n > 30$, pode usar-se a aproximação à distribuição normal. Em tais casos, $\mu = 0$ e $Var = n/(n-2)$.
- À medida que n aumenta, a distribuição tende para a distribuição normal. Para n grande, a distribuição t-Student tende a ser muito semelhante à distribuição normal.

Distribuição Qui-quadrado, Distribuição t-Student e Distribuição F-Snedecor - Distribuições - Studocu

# Distribuição t-Student: Resumo...

propriedades básicas desta distribuição

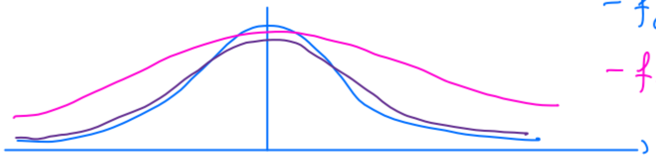
i) a sua função densidade de probabilidade é simétrica em torno de zero e tem uma representação gráfica muito parecida com a da $\mathcal{N}(0,1)$:




$t_N$ lê-se "distribuição t de Student com N graus de liberdade"

ii) Quando n é "bastante grande",

$$f_{t_N}(x) \underset{N \to \infty}{\to} f_{\mathcal{N}(0,1)}(x)$$



# Distribuição t-Student

Formulário

- ***t*-“STUDENT”**

$T = \dfrac{U}{\sqrt{V/n}} \sim t(n)$ com $U \sim N(0,1)$ e $V \sim \chi^2(n)$ independentes

$E(T) = 0$ ; $\mathrm{Var}(T) = \dfrac{n}{n-2}$ $(n > 2)$ ; $\gamma_1 = 0$ ; $\gamma_2 = \dfrac{3(n-2)}{n-4}$ $(n > 4)$

Propriedade:

- Sendo $T \sim t(n) \Rightarrow \lim\limits_{n\to\infty} F_T(t \mid n) = \Phi(t)$

# Distribuição do t-Student: Exercícios

Variáveis Aleatórias Contínuas

2

Suponha que $X \sim t_{(12)}$.

a) Calcule $P[X \leq 2{,}7]$;

b) Qual o valor de $\underline{a}$ tal que $P[X \geq a] = 0{,}95$;

c) Qual o valor de $\underline{b}$ tal que $P[X > b] = 0{,}05$;

d) Qual o valor de $\underline{c}$ tal que $P[-c < X < c] = 0{,}9$.

# Exercício a)

$$t_{n,\varepsilon} : P(X > t_{n,\varepsilon}) = \varepsilon$$

| n \ ε | .400 | .250 | .100 | .050 | .025 | .010 | .005 | .001 |
|---|---|---|---|---|---|---|---|---|
| 1 | .325 | 1,000 | 3.078 | 6.314 | 12.706 | 31.821 | 63.656 | 318.289 |
| 2 | .289 | .816 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 | 22.328 |
| 3 | .277 | .765 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 | 10.214 |
| 4 | .271 | .741 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 | 7.173 |
| 5 | .267 | .727 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 | 5.894 |
| 6 | .265 | .718 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 | 5.208 |
| 7 | .263 | .711 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 4.785 |
| 8 | .262 | .706 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 | 4.501 |
| 9 | .261 | .703 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 | 4.297 |
| 10 | .260 | .700 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 | 4.144 |
| 11 | .260 | .697 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 | 4.025 |
| 12 | .259 | .695 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 | 3.930 |
| 13 | .259 | .694 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 | 3.852 |
| 14 | .258 | .692 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 | 3.787 |
| 15 | .258 | .691 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 | 3.733 |

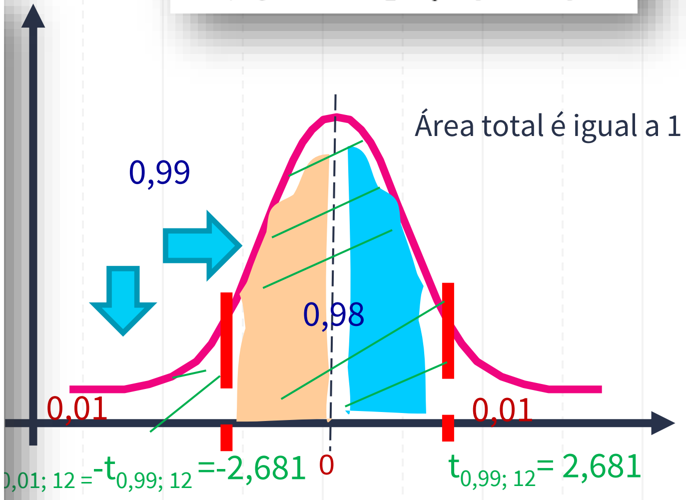
Suponha que X ~ $t_{(12)}$.

a) Calcule P[X ≤ 2,7];

b) Qual o valor de a tal que P[X ≥ a] = 0,95;

c) Qual o valor de b tal que P[X > b] = 0,05;

d) Qual o valor de c tal que P[-c < X < c] = 0,9.




$$P(X \leq 2,7) \sim P(X \leq 2,681) = 1 - P(X > 2,681) = 1 - 0,01 = 0,99$$

# Exercício b)

$$t_{n,\varepsilon} : P(X > t_{n,\varepsilon}) = \varepsilon$$

| ε / n | .400 | .250 | .100 | .050 | .025 | .010 | .005 | .001 |
|---|---|---|---|---|---|---|---|---|
| **1** | .325 | 1,000 | 3.078 | 6.314 | 12.706 | 31.821 | 63.656 | 318.289 |
| **2** | .289 | .816 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 | 22.328 |
| **3** | .277 | .765 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 | 10.214 |
| **4** | .271 | .741 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 | 7.173 |
| **5** | .267 | .727 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 | 5.894 |
| **6** | .265 | .718 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 | 5.208 |
| **7** | .263 | .711 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 4.785 |
| **8** | .262 | .706 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 | 4.501 |
| **9** | .261 | .703 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 | 4.297 |
| **10** | .260 | .700 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 | 4.144 |
| **11** | .260 | .697 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 | 4.025 |
| **12** | .259 | .695 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 | 3.930 |
| **13** | .259 | .694 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 | 3.852 |
| **14** | .258 | .692 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 | 3.787 |
| **15** | .258 | .691 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 | 3.733 |

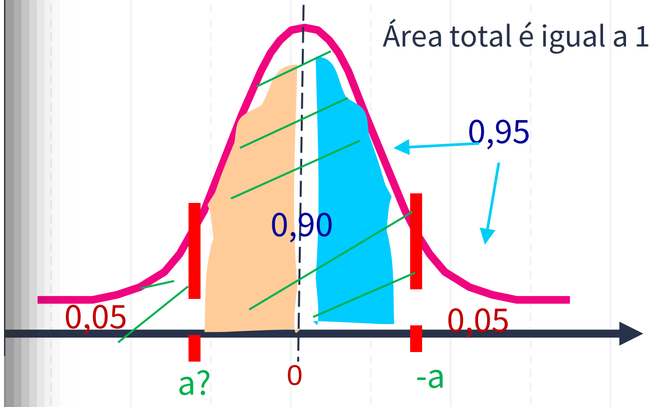
Suponha que $X \sim t_{(12)}$.

a) Calcule $P[X \leq 2{,}7]$;

b) Qual o valor de a tal que $P[X \geq a] = 0{,}95$;

c) Qual o valor de b tal que $P[X > b] = 0{,}05$;

d) Qual o valor de c tal que $P[-c < X < c] = 0{,}9$.




$P(X \geq a) = 0{,}95 \Leftrightarrow P(X>-a) = 0{,}05 \Leftrightarrow -a = 1{,}782 \Leftrightarrow a = -1{,}782$

# Exercício c)

Suponha que X ~ t $_{(12)}$.

a) Calcule P[X ≤ 2,7];

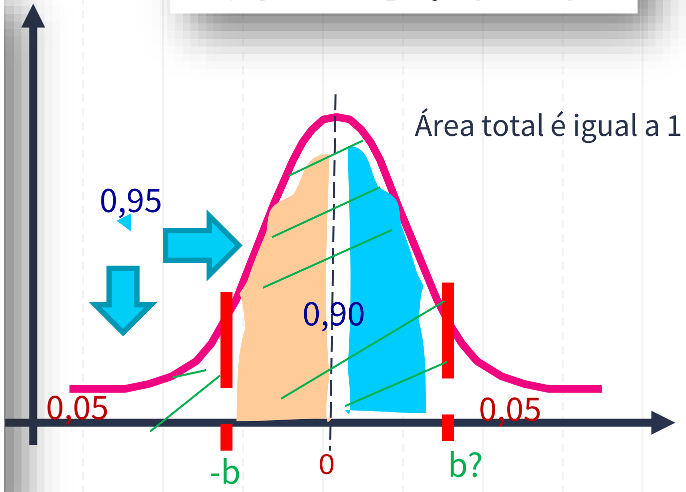
b) Qual o valor de a tal que P[X ≥ a] = 0,95;

c) Qual o valor de b tal que P[X > b] = 0,05;

d) Qual o valor de c tal que P[-c < X < c] = 0,9.

$$t_{n,\varepsilon} : P(X > t_{n,\varepsilon}) = \varepsilon$$

| n \ ε | .400 | .250 | .100 | .050 | .025 | .010 | .005 | .001 |
|---|---|---|---|---|---|---|---|---|
| 1 | .325 | 1,000 | 3.078 | 6.314 | 12.706 | 31.821 | 63.656 | 318.289 |
| 2 | .289 | .816 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 | 22.328 |
| 3 | .277 | .765 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 | 10.214 |
| 4 | .271 | .741 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 | 7.173 |
| 5 | .267 | .727 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 | 5.894 |
| 6 | .265 | .718 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 | 5.208 |
| 7 | .263 | .711 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 4.785 |
| 8 | .262 | .706 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 | 4.501 |
| 9 | .261 | .703 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 | 4.297 |
| 10 | .260 | .700 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 | 4.144 |
| 11 | .260 | .697 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 | 4.025 |
| 12 | .259 | .695 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 | 3.930 |
| 13 | .259 | .694 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 | 3.852 |
| 14 | .258 | .692 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 | 3.787 |
| 15 | .258 | .691 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 | 3.733 |




P(X > b) = 0,05 ⇔ b = 1,782

# Exercício d)

Suponha que X ~ t $_{(12)}$.

a) Calcule P[X ≤ 2,7];

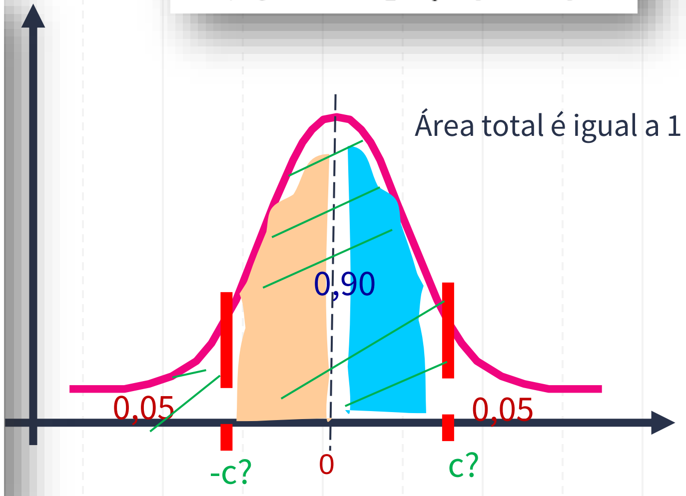
b) Qual o valor de a tal que P[X ≥ a] = 0,95;

c) Qual o valor de b tal que P[X > b] = 0,05;

d) Qual o valor de c tal que P[-c < X < c] = 0,9.

$$t_{n,\varepsilon} : P(X > t_{n,\varepsilon}) = \varepsilon$$

| n \ ε | .400 | .250 | .100 | .050 | .025 | .010 | .005 | .001 |
|---|---|---|---|---|---|---|---|---|
| 1 | .325 | 1,000 | 3.078 | 6.314 | 12.706 | 31.821 | 63.656 | 318.289 |
| 2 | .289 | .816 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 | 22.328 |
| 3 | .277 | .765 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 | 10.214 |
| 4 | .271 | .741 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 | 7.173 |
| 5 | .267 | .727 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 | 5.894 |
| 6 | .265 | .718 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 | 5.208 |
| 7 | .263 | .711 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 4.785 |
| 8 | .262 | .706 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 | 4.501 |
| 9 | .261 | .703 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 | 4.297 |
| 10 | .260 | .700 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 | 4.144 |
| 11 | .260 | .697 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 | 4.025 |
| 12 | .259 | .695 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 | 3.930 |
| 13 | .259 | .694 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 | 3.852 |
| 14 | .258 | .692 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 | 3.787 |
| 15 | .258 | .691 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 | 3.733 |




P(-c < X < c) = 0,90 ⇔ P(X < c) - P(X <-c) = F(c) – F(-c) = F(c) – (1- F(c))
= 2×F(c) – 1 = 0,90 ⇔ F(c) = 0,95 ⇔ c = F(0,95)$^{-1}$ = 1,782

# 3

# Distribuição F-Snedcor

## Variáveis Aleatórias Contínuas

# Distribuição F-Snedcor

- A distribuição F-Snedecor é uma distribuição de probabilidade contínua.
- A distribuição F-Snedecor é dada pelo quociente entre duas variáveis aleatórias com distribuição do qui-quadrado, cada uma dividida pelos respectivos graus de liberdade.

  I.e., se $X \sim \chi^2_{(m)}$ e $Y \sim \chi^2_{(n)}$, duas variáveis aleatórias independentes, então:

$$\boxed{F \equiv \frac{X/m}{Y/n} \cap F_{(m,n)}}$$

- Se X tem uma distribuição F-Snedecor com m e n graus de liberdade, escreve-se:

$$\boxed{X \sim F_{(m,\, n)}}$$

- A distribuição F-Snedecor tem dois parâmetros: m e n – *i.e.*, o nº de graus de liberdade do numerador e do denominador (m, n $\in$ N).

Distribuição Qui-quadrado, Distribuição t-Student e Distribuição F-Snedecor - Distribuições - Studocu

# Distribuição F-Snedcor

- É uma distribuição positiva e não simétrica.
- $E[X] = n/(n-2)$, quando $n > 2$

$$VAR[X] = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)} \quad , \text{ quando } n > 4$$

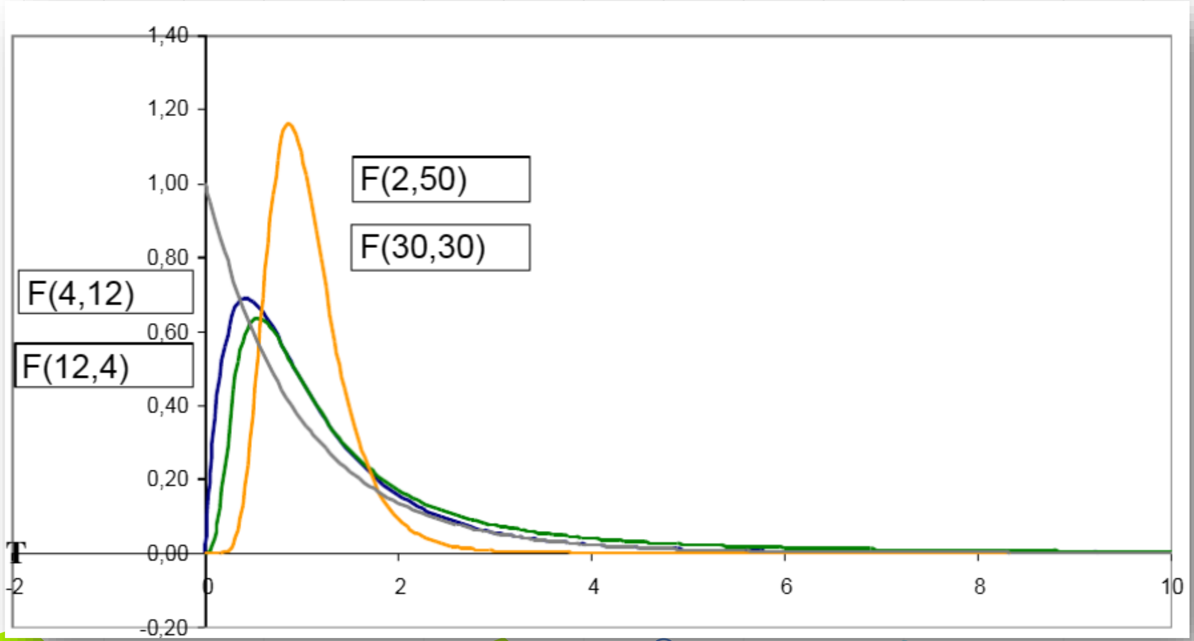
- Se $X \sim F_{(m,\, n)}$, então:

$$\frac{1}{X} \cap F_{(n,m)}$$

- Se $X \sim t_{(n)}$, então $X^2 \sim F_{(1,\, n)}$
- A distribuição F-Snedecor está tabelada para algumas probabilidade e alguns m e n

(m, n = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 20, 30, 60, 120, ∞)

# Distribuição F-Snedcor




Distribuição Qui-quadrado, Distribuição t-Student e Distribuição F-Snedecor - Distribuições - Studocu

# Distribuição F-Snedcor

Formulário

- ***F*-SNEDCOR**

$$F = \frac{U/m}{V/n} \sim F(m,n) \text{ com } U \sim \chi^2(m),\ V \sim \chi^2(n) \text{ (independentes)}$$

$$E(X) = \frac{n}{n-2} \quad (n > 2);\ \mathrm{Var}(X) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)} \quad (n > 4)$$

Propriedades:

- $X \sim F(m,n) \Rightarrow \frac{1}{X} \sim F(n,m)$
- $T \sim t_{(n)} \Rightarrow T^2 \sim F(1,n)$

# Distribuição do F-Snedcor: Exercícios

# 4

Variáveis Aleatórias Contínuas

Suponha que $X \sim F_{(10;\, 5)}$.

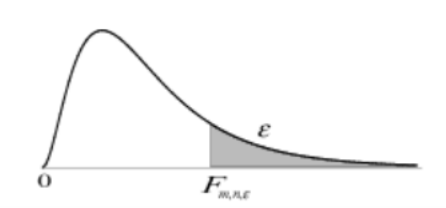
a) Calcule o valor de a tal que $P[X > a] = 0{,}025$;

b) Calcule o valor de b tal que $P[X < b] = 0{,}05$;

c) Calcule os valores de c e d tal que $P[c < X < d] = 0{,}9$.

# Exercício a)

$$F_{m,n,\varepsilon} : P(X > F_{m,n,\varepsilon}) = \varepsilon$$




Suponha que X ~ F $_{(10;\,5)}$.

a) Calcule o valor de <u>a</u> tal que P[X > a] = 0,025;

b) Calcule o valor de <u>b</u> tal que P[X < b] = 0,05;

c) Calcule os valores de <u>c</u> e <u>d</u> tal que P[c < X < d] = 0,9.

m - graus de libe[illegible] do numerador

lib[illegible]de do denominador

| n | ε | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 12 | 15 | 20 | 24 | 30 | 40 | 60 | 120 | ∞ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | .100 | 39.86 | 49.50 | 53.59 | 55.83 | 57.24 | 58.20 | 58.91 | 59.44 | 59.86 | 60.19 | 60.71 | 61.22 | 61.74 | 62.00 | 62.26 | 62.53 | 62.79 | 63.06 | 63.33 |
| | .050 | 161.45 | 199.50 | 215.71 | 224.58 | 230.16 | 233.99 | 236.77 | 238.88 | 240.54 | 241.88 | 243.90 | 245.95 | 248.02 | 249.05 | 250.10 | 251.14 | 252.20 | 253.25 | 254.32 |
| | .025 | 647.79 | 799.48 | 864.15 | 899.60 | 921.83 | 937.11 | 948.20 | 956.64 | 963.28 | 968.63 | 976.72 | 984.87 | 993.08 | 997.27 | 1001.40 | 1005.60 | 1009.79 | 1014.04 | 1018.26 |
| | .010 | 4052.18 | 4999.34 | 5403.53 | 5624.26 | 5763.96 | 5858.95 | 5928.33 | 5980.95 | 6022.40 | 6055.93 | 6106.68 | 6156.97 | 6208.66 | 6234.27 | 6260.35 | 6286.43 | 6312.97 | 6339.51 | 6365.59 |
| 2 | .100 | 8.53 | 9.00 | 9.16 | 9.24 | 9.29 | 9.33 | 9.35 | 9.37 | 9.38 | 9.39 | 9.41 | 9.42 | 9.44 | 9.45 | 9.46 | 9.47 | 9.47 | 9.48 | 9.49 |
| | .050 | 18.51 | 19.00 | 19.16 | 19.25 | 19.30 | 19.33 | 19.35 | 19.37 | 19.38 | 19.40 | 19.41 | 19.43 | 19.45 | 19.45 | 19.46 | 19.47 | 19.48 | 19.49 | 19.50 |
| | .025 | 38.51 | 39.00 | 39.17 | 39.25 | 39.30 | 39.33 | 39.36 | 39.37 | 39.39 | 39.40 | 39.41 | 39.43 | 39.45 | 39.46 | 39.46 | 39.47 | 39.48 | 39.49 | 39.50 |
| | .010 | 98.50 | 99.00 | 99.16 | 99.25 | 99.30 | 99.33 | 99.36 | 99.38 | 99.39 | 99.40 | 99.42 | 99.43 | 99.45 | 99.46 | 99.47 | 99.48 | 99.48 | 99.49 | 99.50 |
| 3 | .100 | 5.54 | 5.46 | 5.39 | 5.34 | 5.31 | 5.28 | 5.27 | 5.25 | 5.24 | 5.23 | 5.22 | 5.20 | 5.18 | 5.18 | 5.17 | 5.16 | 5.15 | 5.14 | 5.13 |
| | .050 | 10.13 | 9.55 | 9.28 | 9.12 | 9.01 | 8.94 | 8.89 | 8.85 | 8.81 | 8.79 | 8.74 | 8.70 | 8.66 | 8.64 | 8.62 | 8.59 | 8.57 | 8.55 | 8.53 |
| | .025 | 17.44 | 16.04 | 15.44 | 15.10 | 14.88 | 14.73 | 14.62 | 14.54 | 14.47 | 14.42 | 14.34 | 14.25 | 14.17 | 14.12 | 14.08 | 14.04 | 13.99 | 13.95 | 13.90 |
| | .010 | 34.12 | 30.82 | 29.46 | 28.71 | 28.24 | 27.91 | 27.67 | 27.49 | 27.34 | 27.23 | 27.05 | 26.87 | 26.69 | 26.60 | 26.50 | 26.41 | 26.32 | 26.22 | 26.13 |
| 4 | .100 | 4.54 | 4.32 | 4.19 | 4.11 | 4.05 | 4.01 | 3.98 | 3.95 | 3.94 | 3.92 | 3.90 | 3.87 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| | .050 | 7.71 | 6.94 | 6.59 | 6.39 | 6.26 | 6.16 | 6.09 | 6.04 | 6.00 | 5.96 | 5.91 | 5.86 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| | .025 | 12.22 | 10.65 | 9.98 | 9.60 | 9.36 | 9.20 | 9.07 | 8.98 | 8.90 | 8.84 | 8.75 | 8.66 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| | .010 | 21.20 | 18.00 | 16.69 | 15.98 | 15.52 | 15.21 | 14.98 | 14.80 | 14.66 | 14.55 | 14.37 | 14.20 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| 5 | .100 | 4.06 | 3.78 | 3.62 | 3.52 | 3.45 | 3.40 | 3.37 | 3.34 | 3.32 | 3.30 | 3.27 | 3.24 | 3.21 | 3.19 | 3.17 | 3.16 | 3.14 | 3.12 | 3.11 |
| | .050 | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.88 | 4.82 | 4.77 | 4.74 | 4.68 | 4.62 | 4.56 | 4.53 | 4.50 | 4.46 | 4.43 | 4.40 | 4.37 |
| | .025 | 10.01 | 8.43 | 7.76 | 7.39 | 7.15 | 6.98 | 6.85 | 6.76 | 6.68 | 6.62 | 6.52 | 6.43 | 6.33 | 6.28 | 6.23 | 6.18 | 6.12 | 6.07 | 6.02 |
| | .010 | 16.26 | 13.27 | 12.06 | 11.39 | 10.97 | 10.67 | 10.46 | 10.29 | 10.16 | 10.05 | 9.89 | 9.72 | 9.55 | 9.47 | 9.38 | 9.29 | 9.20 | 9.11 | 9.02 |

P(X > a) = 0,025 => a = 6,62

# Exercício b)

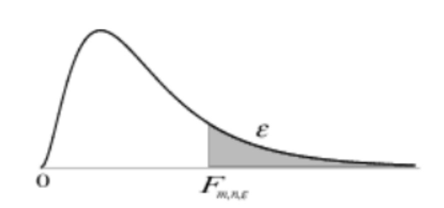
Suponha que X ~ F $_{(10;\,5)}$.

a) Calcule o valor de a tal que P[X > a] = 0,025;

b) Calcule o valor de b tal que P[X < b] = 0,05;

c) Calcule os valores de c e d tal que P[c < X < d] = 0,9.

$$F_{m,n,\varepsilon} : P(X > F_{m,n,\varepsilon}) = \varepsilon$$




| | | m – graus de liberdade do numerador | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ε | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 12 | 15 | 20 | 24 | 30 | 40 | 60 | 120 | ∞ |
| 10 | .100 | 3.29 | 2.92 | 2.73 | 2.61 | 2.52 | 2.46 | 2.41 | 2.38 | 2.35 | 2.32 | 2.28 | 2.24 | 2.20 | 2.18 | 2.16 | 2.13 | 2.11 | 2.08 | 2.06 |
| | .050 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.14 | 3.07 | 3.02 | 2.98 | 2.91 | 2.85 | 2.77 | 2.74 | 2.70 | 2.66 | 2.62 | 2.58 | 2.54 |
| | .025 | 6.94 | 5.46 | 4.83 | 4.47 | 4.24 | 4.07 | 3.95 | 3.85 | 3.78 | 3.72 | 3.62 | 3.52 | 3.42 | 3.37 | 3.31 | 3.26 | 3.20 | 3.14 | 3.08 |
| | .010 | 10.04 | 7.56 | 6.55 | 5.99 | 5.64 | 5.39 | 5.20 | 5.06 | 4.94 | 4.85 | 4.71 | 4.56 | 4.41 | 4.33 | 4.25 | 4.17 | 4.08 | 4.00 | 3.91 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 2.54 | 2.45 | 2.39 | 2.34 | 2.30 | 2.27 | 2.25 | 2.21 | 2.17 | 2.12 | 2.10 | 2.08 | 2.05 | 2.03 | 2.00 | 1.97 |
| | [illegible] | [illegible] | [illegible] | [illegible] | 3.36 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 2.61 | 2.57 | 2.53 | 2.49 | 2.45 | 2.40 |
| | [illegible] | [illegible] | [illegible] | [illegible] | 4.28 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 3.17 | 3.12 | 3.06 | 3.00 | 2.94 | 2.88 |
| | [illegible] | [illegible] | [illegible] | [illegible] | 5.67 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 4.02 | 3.94 | 3.86 | 3.78 | 3.69 | 3.60 |
| [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 2.48 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 2.04 | 2.01 | 1.99 | 1.96 | 1.93 | 1.90 |
| | [illegible] | [illegible] | [illegible] | [illegible] | 3.26 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 2.51 | 2.47 | 2.43 | 2.38 | 2.34 | 2.30 |
| | .025 | 6.55 | 5.10 | 4.47 | 4.12 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 3.02 | 2.96 | 2.91 | 2.85 | 2.79 | 2.72 |
| | .010 | 9.33 | 6.93 | 5.95 | 5.41 | 5.06 | 4.82 | 4.64 | 4.50 | 4.39 | 4.30 | 4.16 | 4.01 | 3.86 | 3.78 | 3.70 | 3.62 | 3.54 | 3.45 | 3.36 |

Quantil da distribuição F-Snedecor de probabilidade 0,05 com 10 e 5 graus de liberdade

Quantil da distribuição F-Snedecor de probabilidade 0,95 com 5 e 10 graus de liberdade. No caso da tabela é o valor d tal que P(F > d) = 0,05

P(X < b) = 0,05 ⇔ 1-P(X > b) = 0,05 ⇔ P(X > b) = 0,95

=> b = F0,05; 10,5 = 1/ F0,95; 5,10 = 1/3,33 = 0,30

$$F_{n_1-1,n_2-1;\frac{\alpha}{2}} = \frac{1}{F_{n_2-1,n_1-1;1-\frac{\alpha}{2}}}$$

# Exercício c)

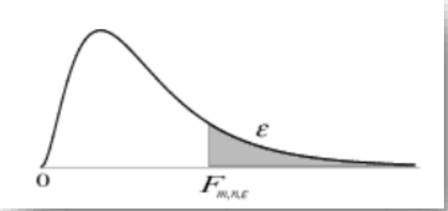
Suponha que X ~ F $_{(10;\,5)}$.

a) Calcule o valor de a tal que P[X > a] = 0,025;

b) Calcule o valor de b tal que P[X < b] = 0,05;

c) Calcule os valores de c e d tal que P[c < X < d] = 0,9.

$$F_{m,n,\varepsilon} : P(X > F_{m,n,\varepsilon}) = \varepsilon$$




| | | m - graus de liberdade do numerador | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ε | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 12 | 15 | | | | | | | ∞ |
| 1 | .100 | 39.86 | 49.50 | 53.59 | 55.83 | 57.24 | 58.20 | 58.91 | 59.44 | 59.86 | 60.19 | 60.71 | 61.2 | | | | | | | 63.33 |
| | .050 | 161.45 | 199.50 | 215.71 | 224.58 | 230.16 | 233.99 | 236.77 | 238.88 | 240.54 | 241.88 | 243.90 | 245.9 | | | | | | | 254.32 |
| | .025 | 647.79 | 799.48 | 864.15 | 899.60 | 921.83 | 937.11 | 948.20 | 956.64 | 963.28 | 968.63 | 976.72 | 984.8 | | | | | | | 1018.26 |
| | .010 | 4052.18 | 4999.34 | 5403.53 | 5624.26 | 5763.96 | 5858.95 | 5928.33 | 5980.95 | 6022.40 | 6055.93 | 6106.68 | 6156.9 | | | | | | | 6365.59 |
| 2 | .100 | 8.53 | 9.00 | 9.16 | 9.24 | 9.29 | 9.33 | 9.35 | 9.37 | 9.38 | 9.39 | 9.41 | 9.42 | 9.44 | 9.45 | 9.46 | 9.47 | 9.47 | 9.48 | 9.49 |
| | .050 | 18.51 | 19.00 | 19.16 | 19.25 | 19.30 | 19.33 | 19.35 | 19.37 | 19.38 | 19.40 | 19.41 | 19.43 | 19.45 | 19.45 | 19.4 | | | | |
| | .025 | 38.51 | 39.00 | 39.17 | 39.25 | 39.30 | | | | | | | | | | | | | | |
| | .010 | 98.50 | 99.00 | 99.16 | 99.25 | 99.30 | | | | | | | | | | | | | | |
| 3 | .100 | 5.54 | 5.46 | 5.39 | 5.34 | 5.31 | | | | | | | | | | | | | | |
| | .050 | 10.13 | 9.55 | 9.28 | 9.12 | 9.01 | | | | | | | | | | | | | | |
| | .025 | 17.44 | 16.04 | 15.44 | 15.10 | 14.88 | | | | | | | | | | | | | | |
| | .010 | 34.12 | 30.82 | 29.46 | 28.71 | 28.24 | | | | | | | | | | | | | | |
| 4 | .100 | 4.54 | 4.32 | 4.19 | 4.11 | 4.05 | | | | | | | | | | | | | | |
| | .050 | 7.71 | 6.94 | 6.59 | 6.39 | 6.26 | | | | | | | | | | | | | | |
| | .025 | 12.22 | 10.65 | 9.98 | 9.60 | 9.36 | | | | | | | | | | | | | | |
| | .010 | 21.20 | 18.00 | 16.69 | 15.98 | 15.52 | | | | | | | | | | | | | | |
| 5 | .100 | 4.06 | 3.78 | 3.62 | 3.52 | 3.45 | 3.40 | 3.37 | 3.34 | 3.32 | 3.30 | 3.27 | 3.24 | 3.21 | 3.19 | 3.17 | 3.16 | 3.14 | 3.12 | 3.11 |
| | .050 | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.88 | 4.82 | 4.77 | 4.74 | 4.68 | 4.62 | 4.56 | 4.53 | 4.50 | 4.46 | 4.43 | 4.40 | 4.37 |
| | .025 | 10.01 | 8.43 | 7.76 | 7.39 | 7.15 | 6.98 | 6.85 | 6.76 | 6.68 | 6.62 | 6.52 | 6.43 | 6.33 | 6.28 | 6.23 | 6.18 | 6.12 | 6.07 | 6.02 |
| | .010 | 16.26 | 13.27 | 12.06 | 11.39 | 10.97 | 10.67 | 10.46 | 10.29 | 10.16 | 10.05 | 9.89 | 9.72 | 9.55 | 9.47 | 9.38 | 9.29 | 9.20 | 9.11 | 9.02 |

liberdade do denominador

Quantil da distribuição F-Snedecor de probabilidade 0,05 com 10 e 5 graus de liberdade

$$F_{n_1-1,n_2-1;\frac{\alpha}{2}} = \frac{1}{F_{n_2-1,n_1-1;1-\frac{\alpha}{2}}}$$

P(c < X < d) = 0,9
P(X > d) = 0,05 => d = 4,74
P(X > c) = 0,95 => c = F0,05; 10,5 = 1/ F0,95; 5,10 = 1/3,33 = 0,30
(ver slide a seguir)

# Exercício c)

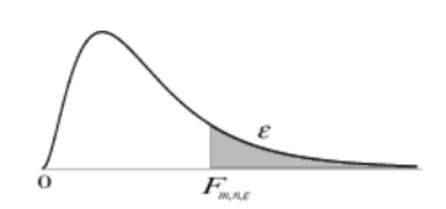
Suponha que X ~ F $_{(10;\,5)}$.

a) Calcule o valor de a tal que P[X > a] = 0,025;

b) Calcule o valor de b tal que P[X < b] = 0,05;

c) Calcule os valores de c e d tal que P[c < X < d] = 0,9.

$$F_{m,n,\varepsilon} : P(X > F_{m,n,\varepsilon}) = \varepsilon$$




| | | m – graus de liberdade do numerador | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | ε | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 12 | 15 | 20 | 24 | 30 | 40 | 60 | 120 | ∞ |
| 10 | .100 | 3.29 | 2.92 | 2.73 | 2.61 | 2.52 | 2.46 | 2.41 | 2.38 | 2.35 | 2.32 | 2.28 | 2.24 | 2.20 | 2.18 | 2.16 | 2.13 | 2.11 | 2.08 | 2.06 |
| | .050 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.14 | 3.07 | 3.02 | 2.98 | 2.91 | 2.85 | 2.77 | 2.74 | 2.70 | 2.66 | 2.62 | 2.58 | 2.54 |
| | .025 | 6.94 | 5.46 | 4.83 | 4.47 | 4.24 | 4.07 | 3.95 | 3.85 | 3.78 | 3.72 | 3.62 | 3.52 | 3.42 | 3.37 | 3.31 | 3.26 | 3.20 | 3.14 | 3.08 |
| | .010 | 10.04 | 7.56 | 6.55 | 5.99 | 5.64 | 5.39 | 5.20 | 5.06 | 4.94 | 4.85 | 4.71 | 4.56 | 4.41 | 4.33 | 4.25 | 4.17 | 4.08 | 4.00 | 3.91 |
| 11 | .100 | 3.23 | 2.86 | 2.66 | 2.54 | 2.45 | 2.39 | 2.34 | 2.30 | 2.27 | 2.25 | 2.21 | 2.17 | 2.12 | 2.10 | 2.08 | 2.05 | 2.03 | 2.00 | 1.97 |
| | .050 | 4.84 | 3.98 | 3.59 | 3.36 | 3.20 | 3.09 | 3.01 | 2.95 | 2.90 | 2.85 | 2.79 | 2.72 | 2.65 | 2.61 | 2.57 | 2.53 | 2.49 | 2.45 | 2.40 |
| | .025 | 6.72 | 5.26 | 4.63 | 4.28 | 4.04 | 3.88 | 3.76 | 3.66 | 3.59 | 3.53 | 3.43 | 3.33 | 3.23 | 3.17 | 3.12 | 3.06 | 3.00 | 2.94 | 2.88 |
| | .010 | 9.65 | 7.21 | 6.22 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 3.60 |
| 12 | .100 | 3.18 | 2.81 | 2.61 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 1.90 |
| | .050 | 4.75 | 3.89 | 3.49 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 2.30 |
| | .025 | 6.55 | 5.10 | 4.47 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 2.72 |
| | .010 | 9.33 | 6.93 | 5.95 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | 3.36 |

P(c < X < d) = 0,9

P(X > c) = 0,95 => c = F0,05; 10,5 = 1/ F0,95; 5,10 = 1/3,33 = 0,30

$$F_{n_1-1,n_2-1\,;\frac{\alpha}{2}} = \frac{1}{F_{n_2-1,n_1-1\,;1-\frac{\alpha}{2}}}$$

# 5 Distribuição Gama

Variáveis Aleatórias Contínuas

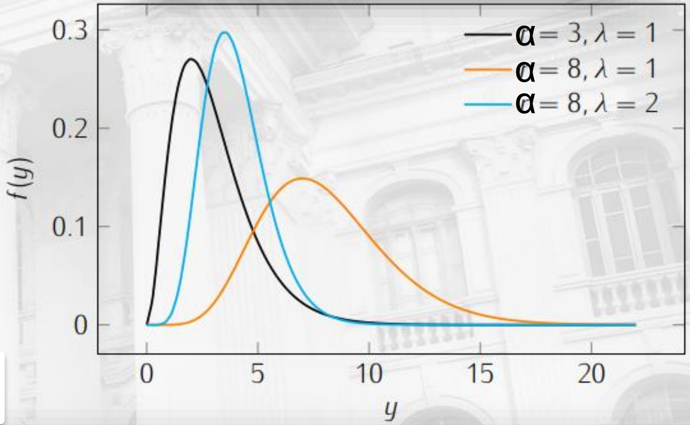
A distribuição exponencial prevê o tempo de espera até o **primeiro** evento. A distribuição gama, por outro lado, prevê o tempo de espera até que o evento **k-ésimo evento** ocorra.

# Distribuição Gama

Distribuições contínuas: Lognormal, Gama, Weibull e Beta (ufpr.br)

- Seja $Y_{Ei} \sim \text{Exp}(\lambda)$ $(i \in \{1, \ldots, k\})$ uma variável com distribuição Exponencial. Então, $Y = Y_{E1} + Y_{E2} + \cdots + Y_{Ek}$ tem distribuição Gama.
- A Gama tem suporte no conjunto dos **reais positivos**, assumindo formas assimétricas.
- Ela tem aplicações na área de **confiabilidade** e análise de **sobrevivência**, assim como a Lognormal.

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




Formulário

- **GAMA** $X \sim G(\alpha, \lambda)$, $(\lambda > 0, \alpha > 0)$

$$f(x \mid \alpha, \lambda) = \frac{\lambda^{\alpha} e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)}, \quad x > 0; \; E(X) = \frac{\alpha}{\lambda}; \; \text{Var}(X) = \frac{\alpha}{\lambda^2}; \; M_X(s) = \left(\frac{\lambda}{\lambda - s}\right)^{\alpha}, s < \lambda; \; \gamma_1 = \frac{2}{\sqrt{\alpha}}; \; \gamma_2 = 3 + \frac{6}{\alpha}$$

Propriedades:

- $X_i \sim G(\alpha_i; \lambda)$, $(i = 1, 2, \ldots, k)$ independentes $\Rightarrow \sum_{i=1}^{k} X_i \sim G\left(\sum_{i=1}^{k} \alpha_i; \lambda\right)$
- $X \sim G(\alpha, \lambda)$ então $Y = cX \sim G(\alpha, \frac{\lambda}{c})$ onde $c$ constante positiva

# Distribuição Gama: Exemplos

1. Soma de v.a. com distribuição Exponencial.
2. Tempo de carregamento de um navio.
3. Volume de chuva em dias com precipitação.
4. Tempo de permanência de um usuário em um site.
5. Distribuição de idade de animais em ambiente natural.
6. Tempo de vida de um paciente após transplante.
7. Distância dos passes de bola em um jogo de futebol.

# Distribuição Gama: Função Densidade de Probabilidade

As **distribuições Exponencial** e **Qui-quadrado** são casos particulares de uma distribuição mais geral, a **distribuição Gama**.

Definição: A v. a. contínua X segue uma distribuição Gama com parâmetros $\boldsymbol{\alpha}$ e $\boldsymbol{\lambda}$, i. e., $\boldsymbol{X \sim G(\alpha;\ \lambda)}$, se a sua função densidade de probabilidade é:

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

onde $\Gamma(\alpha)$ é a função Gama, definida por $\Gamma(\alpha) = \int_0^{+\infty} e^{-x}x^{\alpha-1}dx$.

O parâmetros caracterizadores desta distribuição são $\alpha$ e $\lambda$.

ProbabilidadesEstatistica_2019 (uevora.pt)

Distribuições contínuas: Lognormal, Gama, Weibull e Beta (ufpr.br)

# Distribuição Gama: Função Gama

Para cada número positivo $\alpha$, seja $\Gamma(\alpha)$ definido como:

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

- $\Gamma(1) = \int_0^\infty e^{-x}\, dx = 1$
- Se $\alpha > 1$, então $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$
- $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$
- $\Gamma(n) = (n-1)!$, $n$ interiro positivo.

# Distribuição Gama: Relação com outras Distribuições

- A distribuição Gama tem como **caso particular** a distribuição exponencial ($\lambda$) ao fixarmos $\alpha = 1$.
- Dessa relação, a Gama pode ser obtida como o **tempo acumulado** para $k$ eventos de Poisson, uma vez que o intervalo entre eventos é Exponencial.
- A Gama tem mais variedades de formas por ter 2 parâmetros, permitindo modelar adequadamente um maior número de variáveis aleatórias que a Exponencial.
- A **soma** de v.a. Gama é Gama, ou seja, se $Y_1, Y_2, \cdots, Y_k$ são variáveis aleatórias independentes, com distribuição Gama de parâmetros $\alpha$ e $\lambda$, então

$$Y_{\text{soma}} = Y_1 + Y_2 + \cdots + Y_k \sim \text{Gama}(k\alpha, \lambda).$$

- A distribuição Erlang é um **caso particular** da Gama quando $r$ é um número natural, $r \in \{1, 2, \ldots\}$.

Distribuições contínuas: Lognormal, Gama, Weibull e Beta (ufpr.br)

# Distribuição Gama: Relação com outras Distribuições

**Casos particulares:**

- $X \sim \chi_n^2 \Leftrightarrow X \sim G\left(\alpha = \frac{n}{2};\ \lambda = \frac{1}{2}\right);$
- $X \sim Exp(\lambda) \Leftrightarrow X \sim G(\alpha = 1;\ \lambda).$

ProbabilidadesEstatistica_2019 (uevora.pt)

# Distribuição Gama

α Parâmetro de forma
λ Parâmetro de taxa

O aspeto da distribuição depende do valor dos parâmetros (Figura 5.13).

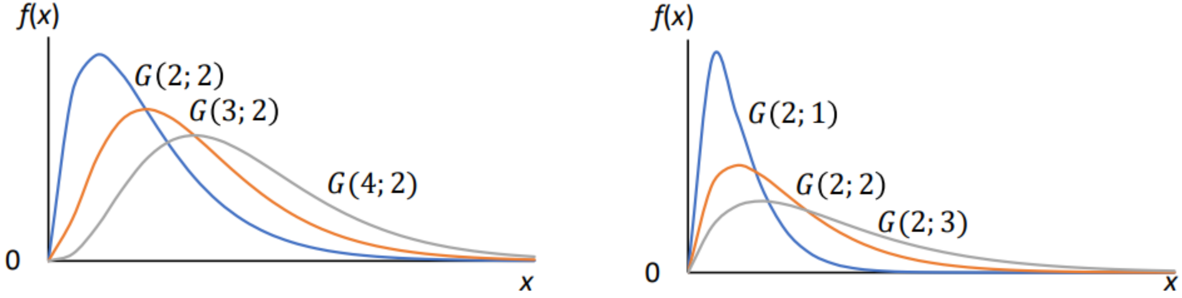



***Figura 5.13:*** *Função densidade de probabilidade da distribuição Gama para diferentes valores de* $\alpha$ *e* $\lambda$.

Se $X \sim G(\alpha;\ \lambda)$, então $\mu_X = E(X) = \frac{\alpha}{\lambda}$ e $\sigma_X^2 = Var(X) = \frac{\alpha}{\lambda^2}$.

**Teorema da aditividade:** Se $X_i, i = 1, 2, \ldots, K$, são v. a. independentes e $X_i \sim G(\alpha_i;\ \lambda)$, então

$$\sum_{i=1}^{K} X_i \sim G(\alpha;\ \lambda), \text{com } \alpha = \sum_{i=1}^{K} \alpha_i.$$

ProbabilidadesEstatistica_2019 (uevora.pt)

# Distribuição Gama

A distribuição Gama pode ser como uma generalização da distribuição Exponencial para descrever a v.a. $X$ que representa **o tempo de espera até à ocorrência do $n$-ésimo sucesso**. A variável $X$ resulta da soma dos

tempos de espera entre as várias ocorrências sucessivas $(X_i)$ até à ocorrência pretendida. Deste modo, pelo teorema da aditividade como $X_i, i = 1, \dots, n$, são v. a. independentes e $X_i \sim Exp(\lambda) \Leftrightarrow X_i \sim G(1;\ \lambda)$, então

$$X = \sum_{i=1}^{n} X_i \sim G(n;\ \lambda).$$

# 6 Teorema do Limite Central

# Distribuição Assintótica da Soma e da Média de V.A.´s: Teorema do Limite Central (TLC)

- Na maioria das situações é difícil determinar a distribuição da soma de variáveis (mesmo que sejam independentes!). O teorema seguinte **justifica a grande utilidade** e **importância da distribuição normal** (quer em probabilidades quer em estatística).

- O teorema que vamos ver de seguida diz-nos que, para um grande número de v.a.'s com idêntico comportamento (distribucional), seja ele qual for, e mesmo que seja praticamente desconhecido, a distribuição da soma das v.a.'s aproxima-se de uma distribuição Normal; e tem uma distribuição que é tão mais próxima da Normal, quanto maior for o número de v.a.'s da soma.

# TLC

**Teorema do Limite Central**

Seja $X_1, \ldots, X_n$ uma sequência de v.a. independentes e identicamente distribuídas com valor esperado $\mu < \infty$ e variância $\sigma^2 < \infty$. Considere-se $S_n = \sum_{i=1}^{n} X_i$ e $\bar{X}_n = \frac{S_n}{n}$, então quando $n \to +\infty$

$$\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \overset{a}{\sim} N(0,1),$$

ou seja,

$$\lim_{n \to +\infty} P\left(\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} \leq z\right) = \Phi(z),$$

onde $\Phi(.)$ é a função de distribuição da normal reduzida i.e., $N(0,1)$.

Slides da Professora Conceição Amado

# TLC

## Teorema do Limite Central

De modo equivalente, se tem

$$\frac{\bar{X}_n - E\left(\bar{X}_n\right)}{\sqrt{V\left(\bar{X}_n\right)}} = \frac{\bar{X}_n - \mu}{\sqrt{\sigma^2/n}} \overset{a}{\sim} N(0,1),$$

ou seja,

$$\lim_{n \to +\infty} P\left(\frac{\bar{X}_n - E\left(\bar{X}_n\right)}{\sqrt{V\left(\bar{X}_n\right)}} \leq z\right) = \Phi(z)$$

Slides da Professora Conceição Amado

# Valor Médio e Variância de Somas e Médias de V.A.´s Independentes

**Observar que:**

- $$E\left(S_n\right) = E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i) = nE(X_1) = n\mu$$

  e como $X_1, X_2 \cdots X_n$ são v.a. independentes tem-se

  $$V\left(S_n\right) = V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) = nV(X_1) = n\sigma^2$$

- $$E\left(\bar{X}_n\right) = E\left(\sum_{i=1}^{n} X_i/n\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = E(X_1) = \mu$$

  e como $X_1, X_2 \cdots X_n$ são v.a. independentes tem-se

  $$V\left(\bar{X}_n\right) = V\left(\sum_{i=1}^{n} X_i/n\right) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n}V(X_1) = \frac{\sigma^2}{n}$$

Formulário

**AMOSTRAGEM. DISTRIBUIÇÕES POR AMOSTRAGEM**

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \; ; \quad S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n} = \frac{\sum_{i=1}^{n} X_i^2}{n} - \bar{X}^2 \; ; \quad (n-1)S'^2 = n\,S^2$$

$$E(\bar{X}) = \mu \; ; \quad \mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n} \; ; \quad E(S^2) = \frac{n-1}{n}\sigma^2 \; ; \quad E(S'^2) = \sigma^2$$

# TLC: Aproximações

**Observações:**

- O TLC deve ser usado apenas quando as v.a.'s $X_1, \ldots, X_n$ não têm distribuição normal!
- A demonstração do teorema exige algumas ferramentas matemáticas avançadas.
- As v.a. $X_1, \ldots, X_n$ podem ser discretas ou contínuas.
- Geralmente considera-se $n$ grande se $n \geq 30$
- As distribuições binomial e de Poisson podem ser aproximadas pela distribuição normal (na secção anterior vimos que podem ser escritas como somas de variáveis aleatórias). Assim:
  - $X \sim Bin(n, p)$ pode ser aproximada por $\tilde{X} \sim N(np, np(1-p))$, a aproximação tem menor erro quando $np > 5$ e $n(1-p) > 5$
  - $X \sim Poisson(\lambda)$ pode ser aproximada por $\tilde{X} \sim N(\lambda, \lambda)$, a aproximação tem menor erro quando $\lambda > 5$

# Teorema de De Moivre-Laplace: Binomial - Normal

**T. Limite Central: Aplicação a aproximações para distribuições discretas**

**Corolário 5.3 (T. de De Moivre-Laplace)** – Dada a sucessão de variáveis aleatórias iid, $X_1, X_2, \ldots, X_n, \ldots$, com distribuição de Bernoulli de média $E(X_i) = \theta$ e, portanto, $\mathrm{Var}(X_i) = \theta(1-\theta)$, tem-se,

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

Nota: $X = \sum_{i=1}^{n} X_i \sim B(n;\theta)$. Quando $n$ é grande, utilizar o corolário. Mas...

**com Correção de Continuidade.**

Murteira et al (2015)

# Aproximações Baseadas no TLC: Binomial - Normal

- Probabilidades associadas a uma distribuição Binomial, B(n,p), podem ser aproximadas utilizando uma distribuição Normal, N($\mu$,$\sigma$), com $\mu$=np e $\sigma = \sqrt{np(1-p)}$.

  Para que a aproximação não seja muito má, devemos ter np $\geq$ 5 e n (1-p) $\geq$ 5.

Aproximação da Normal para a Distribuição Binomial

- Quanto mais perto p estiver de 0,5, melhor será a aproximação da normal para a distribuição binomial
- Quanto maior o tamanho da amostra, n, melhor será a aproximação da normal para a distribuição binomial
- Regra Geral:
  - A distribuição normal pode ser utilizada para aproximar a distribuição binomial se

$$np \geq 5 \text{ e } n(1-p) \geq 5$$

SlidesBioestatística6 (ua.pt)

# Aproximações Baseadas no TLC: Poisson - Normal

- Probabilidades associadas a uma distribuição de Poisson, P(λ), podem ser aproximadas utilizando uma distribuição Normal, N(μ,σ), com $\mu = \lambda$ e $\sigma = \sqrt{\lambda}$.

  A aproximação será tanto melhor quanto maior for λ.

SlidesBioestatística6 (ua.pt)

# TLC: Correção de Continuidade

A variante do T. L. C. para distribuições discretas introduz a correção de continuidade que permite aproximar uma distribuição discreta por uma distribuição contínua, neste caso a distribuição Normal.

**Correção de continuidade:** Seja $X$ uma v. a. discreta, com variação $\Delta$ (i.e., $X$ pode tomar os valores $0, \Delta, 2\Delta, 3\Delta, \ldots$), com $E(X) = \mu_X$ e $Var(X) = \sigma_X^2$. Então a correção de continuidade é efetuada da seguinte forma:

$$P(X = k) \approx P\left(k - \frac{\Delta}{2} \leq X \leq k + \frac{\Delta}{2}\right).$$

ProbabilidadesEstatistica2019.pdf

# TLC: Correção de Continuidade

Note-se que nas distribuições usuais, por exemplo, Binomial e Poisson, $\Delta = 1$, pois tratam-se de distribuições de contagem logo

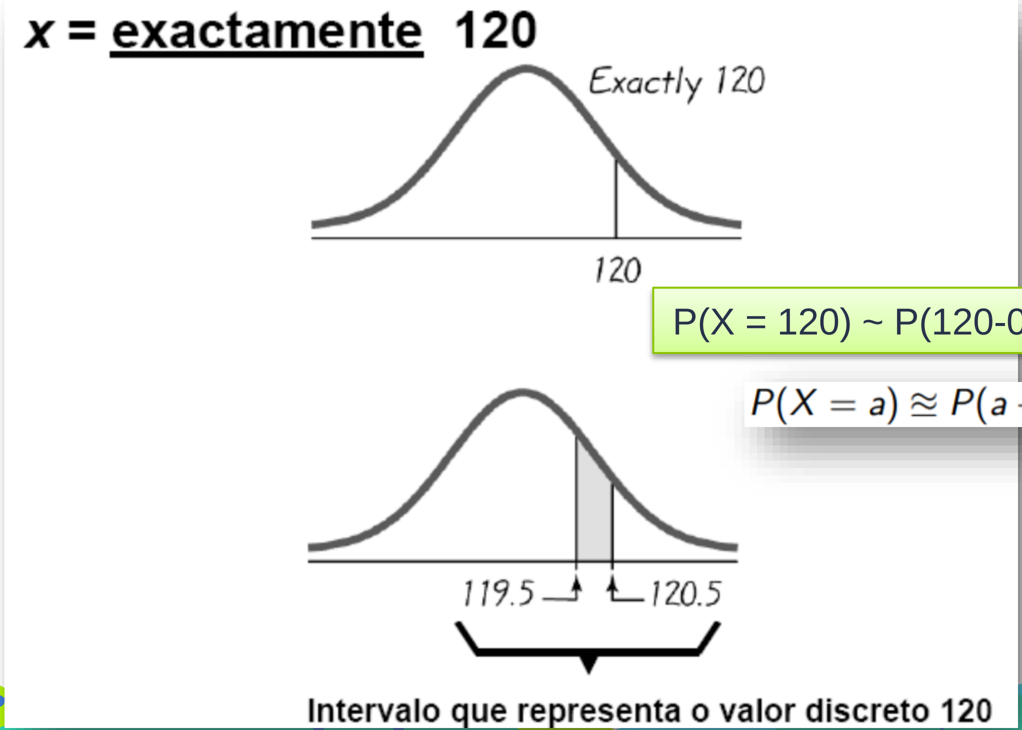
$$P(X = k) \approx P(k - 0{,}5 \leq X < k + 0{,}5).$$

Com a aplicação das propriedades da distribuição Normal, podem generalizar-se as seguintes regras:

- $P(X \leq k) \approx P\left(Z \leq \dfrac{k + 0{,}5 - \mu_X}{\sigma_X}\right) = \Phi\left(\dfrac{k + 0{,}5 - \mu_X}{\sigma_X}\right);$
- $P(X < k) \approx P\left(Z < \dfrac{k - 0{,}5 - \mu_X}{\sigma_X}\right) = \Phi\left(\dfrac{k - 0{,}5 - \mu_X}{\sigma_X}\right);$
- $P(X \geq k) \approx P\left(Z \geq \dfrac{k - 0{,}5 - \mu_X}{\sigma_X}\right) = 1 - \Phi\left(\dfrac{k - 0{,}5 - \mu_X}{\sigma_X}\right);$
- $P(X > k) \approx P\left(Z > \dfrac{k + 0{,}5 - \mu_X}{\sigma_X}\right) = 1 - \Phi\left(\dfrac{k + 0{,}5 - \mu_X}{\sigma_X}\right).$

ProbabilidadesEstatistica2019.pdf

# Aproximações Baseadas no TLC: Correção de Continuidade



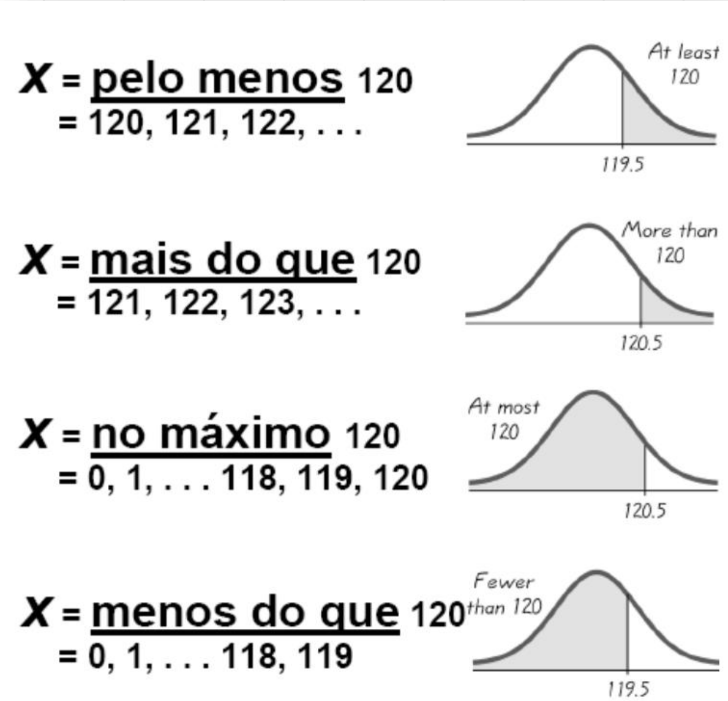

Note-se que $\tilde{X}$ é uma v.a. Normal que “aproxima” X (sendo esta uma v.a. discreta).

P(X = 120) ~ P(120-0,5 < $\tilde{X}$ < 120+0,5)

$$P(X = a) \approxeq P(a - \delta < \tilde{X} < a + \delta)$$

SlidesBioestatística6 (ua.pt)

# Aproximações Baseadas no TLC: Correção de Continuidade




$P(X \geq 120) \sim P(\tilde{X} \geq 120\text{-}0{,}5)$

$P(X > 120) \sim P(\tilde{X} > 120\text{+}0{,}5)$

$P(X \leq 120) \sim P(\tilde{X} \leq 120\text{+}0{,}5)$

$P(X < 120) \sim P(\tilde{X} < 120\text{-}0{,}5)$

Note-se que $\tilde{X}$ é uma v.a. Normal que "aproxima" X (sendo esta uma v.a. discreta).

SlidesBioestatística6 (ua.pt)

# TLC: Resumo da Correção de Continuidade

**Observação - aplicação do TLC para distribuições discretas**

- Importante: A utilização do teorema do limite central para **distribuições discretas** apresenta um problema, uma vez que se vai aproximar um fenómeno discreto por uma distribuição contínua. Embora não exista uma solução ótima para todas as situações, na prática é comum adotar a chamada correção de continuidade. Considere-se $0 < \delta < 1$ e $\tilde{X}$ a v.a. normal que "aproxima" $X$, tem-se:
  - $P(X = a) \approxeq P(a - \delta < \tilde{X} < a + \delta)$
  - $P(a < X < b) \approxeq P(a + \delta \leq \tilde{X} \leq b - \delta)$
  - $P(X \leq a) \approxeq P(\tilde{X} \leq a + \delta)$
  - $P(X < a) \approxeq P(\tilde{X} < a - \delta) = P(\tilde{X} \leq a - \delta)$ (recordar que $\tilde{X}$ é contínua...)
- No caso da aplicação do TLC à binomial e à Poisson o valor típico é $\delta = 0.5$, ou seja metade da variação entre dois valores consecutivos.

Slides da Professora Conceição Amado

# TLC

Formulário

**TEOREMA DO LIMITE CENTRAL E COROLÁRIOS**

TLC: Sendo $X_i$ iid com $E(X_i)=\mu$ e $\mathrm{Var}(X_i)=\sigma^2 \Rightarrow \dfrac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma} = \dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \overset{a}{\sim} N(0,1)$

Corolário: Sendo $X_i \sim B(1;\theta)$, iid $\Rightarrow \dfrac{\sum_{i=1}^{n} X_i - n\theta}{\sqrt{n\theta(1-\theta)}} \overset{a}{\sim} N(0,1)$

Correcção de continuidade: $P(a \le X \le b) \approx \Phi\left(\dfrac{b+\frac{1}{2}-n\theta}{\sqrt{n\theta(1-\theta)}}\right) - \Phi\left(\dfrac{a-\frac{1}{2}-n\theta}{\sqrt{n\theta(1-\theta)}}\right)$, com $a$ e $b$ inteiros

Corolário: Sendo $X \sim \mathrm{Po}(\lambda)$, quando $\lambda \to +\infty \Rightarrow \dfrac{X-\lambda}{\sqrt{\lambda}} \overset{a}{\sim} N(0,1)$

Correcção de continuidade: $P(a \le X \le b) \approx \Phi\left(\dfrac{b+\frac{1}{2}-\lambda}{\sqrt{\lambda}}\right) - \Phi\left(\dfrac{a-\frac{1}{2}-\lambda}{\sqrt{\lambda}}\right)$, com $a$ e $b$ inteiros

# Exemplo 1: TLC

**Exemplo 1- aplicação TLC**

Uma empresa de chocolates embala caixas com 100 pacotes de bombons. Os pesos por pacote são v.a.'s $X_i$, onde $X_i \perp\!\!\!\perp X_j$, com valor médio 0.5kg e variância 0.1kg$^2$. São colocadas 10 caixas numa palete. Qual é a probabilidade aproximada do peso dos bombons da palete ser superior a 510kg?

Slides da Professora Conceição Amado

# Exemplo 1: TLC

**Resolução:**

$X_i -$ 'v.a. peso do $i-$ésimo pacote (kg)'

$E(X_i) = 0.5,\ V(X_i) = 0.1\ ,\forall i$ e $X_i \perp\!\!\!\perp X_j$ para $i \neq j$

$S = \sum_{i=1}^{1000} X_i$ 'v.a. peso dos bombons colocados na palete (kg)'

$$E(S) = E\left(\sum_{i=1}^{1000} X_i\right) = \sum_{i=1}^{1000} E(X_i) = 1000E(X_1) = 1000 \times 0.5 = 500$$

(porque são identicamente distribuídas (i.d.) a $X_1$)

$$V(S) = V\left(\sum_{i=1}^{1000} X_i\right) = \sum_{i=1}^{1000} V(X_i) = 1000V(X_1) = 1000 \times 0.1 = 100$$

porque são independentes e identicamente distribuídas (i.d.) a $X_1$.

Slides da Professora Conceição Amado

# Exemplo 1: TLC

**Resolução (cont.):**

Pelo T.L.C. tem-se que:

$$\frac{S - E(S)}{\sqrt{V(S)}} = \frac{S - 500}{\sqrt{100}} \overset{a}{\sim} N(0,1),$$

$$P(S > 510) = 1 - P(S \leq 510) = 1 - P\left(\frac{S - 500}{\sqrt{100}} \leq \frac{510 - 500}{\sqrt{100}}\right) \underset{T.L.C.}{\approx}$$

$$\underset{T.L.C.}{\approx} 1 - \Phi(1) \approx 0.1587.$$

Slides da Professora Conceição Amado

# Exemplo 2: TLC

**Exemplo 2 - aplicação TLC e correção de continuidade**

O número de chamadas de telemóvel registadas a partir de certa "zona" numa hora tem, em condições estacionárias, distribuição de Poisson de parâmetro 1500. Calcule a probabilidade (aproximada) de ocorrerem mais de 1600 chamadas na próxima hora.

Slides da Professora Conceição Amado

# Exemplo 2: TLC

**Resolução:**

$X \sim Poisson(1500)$, como o valor de $\lambda$ é elevado não consta nas tabelas disponibilizadas, e sem a ajuda de um computador não é possível calcular a probabilidade pedida. Mas, podemos usar a distribuição normal para obter uma aproximação.

Como $X \sim Poisson(1500)$ pode ser aproximada por $\tilde{X} \sim N(1500, 1500)$

$$P(X > 1600) = 1 - P(X \leq 1600) \simeq 1 - P(\tilde{X} \leq 1600 + 0.5) =$$

$$= 1 - P(\tilde{X} \leq 1600.5) = 1 - P\left(\frac{\tilde{X} - 1500}{\sqrt{1500}} \leq \frac{1600.5 - 1500}{\sqrt{1500}}\right) =$$

$$= 1 - \Phi(2.59) \simeq 0.0048$$

**Nota**:
Lambda = 1500 > 5
Logo pode-se aproximar a distribuição Poisson à distribuição Normal

Slides da Professora Conceição Amado

# Teorema do Limite Central: Exercícios

# 7

1. Considere que o tempo de viagem de autocarro entre Lisboa e Évora segue uma distribuição Uniforme entre 100 a 120 minutos. Numa amostra aleatória de 30 viagens, qual a probabilidade de a média dos tempos de viagem ser inferior a 112 minutos?

ProbabilidadesEstatistica2019.pdf

# Exercício 1: TLC

Seja $X_i$ a v.a. que representa o tempo da viagem $i$ de autocarro entre Lisboa e Évora, com $X_i \sim U(100;\ 120)$.
Logo,

$$\mu = E(X_i) = \frac{100 + 120}{2} = 110;$$

$$\sigma^2 = Var(X_i) = \frac{(120 - 100)^2}{12} = \frac{100}{3}.$$

Seja $\overline{X} = \frac{1}{30}(X_1 + \cdots + X_{30})$ a v.a. que representa a média dos 30 tempos de viagem de autocarro entre Lisboa e Évora.
Pelo corolário do T. L. C.,

$$Z = \frac{\overline{X} - 110}{\sqrt{\frac{100}{3 \times 30}}} \overset{\circ}{\sim} N(0; 1).$$

Logo,

$$P(\overline{X} < 112) \approx \mathrm{P}\left(Z < \frac{112 - 110}{\sqrt{\frac{100}{3 \times 30}}}\right) = \Phi(1{,}80) = 0{,}9641.$$

ProbabilidadesEstatistica2019.pdf

2. Seja $X$ a v. a. que representa o tempo decorrido entre a entrada consecutiva de 2 automóveis numa autoestrada (AE), que segue uma distribuição Exponencial com valor médio 15 segundos. Numa amostra aleatória de 100 entradas consecutivas, qual a probabilidade de a soma dos tempos ser superior a 1350 segundos?

ProbabilidadesEstatistica2019.pdf

# Exercício 2: TLC

Seja $X_i$ a v.a. que representa o tempo decorrido entre a $i$-ésima entrada consecutiva de 2 automóveis numa auto-estrada (AE), com $X_i \sim \ Exp\left(\lambda = \frac{1}{15}\right)$. Logo,

$$\mu = E(X_i) = \frac{1}{\frac{1}{15}} = 15 \text{ segundos},$$

$$\sigma^2 = Var(X_i) = \frac{1}{\left(\frac{1}{15}\right)^2} = 15^2 = 225.$$

Seja $S_{100} = X_1 + \cdots + X_{100}$ a v.a. que representa a soma dos 100 decorrido entre a $i$-ésima entrada consecutiva de 2 automóveis numa autoestrada (AE).

Pelo T. L. C.,

$$Z = \frac{S_{100} - 100 \times 15}{\sqrt{100 \times 225}} \overset{\circ}{\sim} N(0;1).$$

Logo,

$$P(S_{100} > 1350) = 1 - P(S_{100} \le 1350) \approx 1 - P\left(Z < \frac{1350 - 1500}{\sqrt{100 \times 225}}\right) = 1 - \Phi(-1) = \Phi(1) = 0{,}8413.$$

ProbabilidadesEstatistica2019.pdf

3. Com a ingestão de um determinado fármaco para o tratamento da depressão, 20% dos doentes referem sentir efeitos secundários durante a 1ª semana de tratamento. Seja $X$ a v. a. que representa o número de doentes que sentem os referidos efeitos secundários numa amostra de $n$ doentes. Numa amostra de 100 doentes, calcule a probabilidade de:

a) Pelo menos 15 terem sentido os efeitos secundários.

b) Mais de 16 e menos de 45 terem sentido efeitos secundários.

ProbabilidadesEstatistica2019.pdf

# Exercício 3: TLC

Sabe-se que $X \sim B(n = 100; p = 0{,}2)$.

Como $n > 50$ e $0{,}1 < p < 0{,}9$ podemos usar a aproximação à distribuição Normal e

$$X \overset{\circ}{\sim} N(\mu = np = 20;\ \sigma = \sqrt{npq} = 4).$$

a) $P(X \geq 15) = 1 - P(X < 15) \approx 1 - P\left(Z < \frac{15 - 0{,}5 - 20}{4}\right) = 1 - \Phi(-1{,}38) = \Phi(1{,}38) = 0{,}9162.$

b) $P(16 < X < 45) \approx P\left(\frac{16 + 0{,}5 - 20}{4} < Z < \frac{45 - 0{,}5 - 20}{4}\right) = P(-0{,}88 < Z < 6{,}13)$

$= \Phi(6{,}13) - \Phi(-0{,}88) = \Phi(6{,}13) - \big(1 - \Phi(0{,}88)\big) \approx 1 - (1 - 0{,}8106) = 0{,}8106.$

ProbabilidadesEstatistica2019.pdf

**Nota**:
n×p = 20 > 5
n×p×(1-p) = 16 > 5
Logo pode-se aproximar a distribuição Binomial à distribuição Normal

4. Seja $X$ a v.a. que representa o número de automóveis que entram numa autoestrada (AE) num período de 30 segundos, com $X \sim P(\lambda = 9)$. Num período de 5 minutos, qual a probabilidade de:

a) Entrarem mais de 59 automóveis na AE?

b) Entrarem no mínimo 55 e no máximo 80 automóveis na AE?

ProbabilidadesEstatistica2019.pdf

# Exercício 4: TLC

**Nota**:
5 minutos = 300 segundos

Para 30 segundos, tem-se lambda = 9, então para 300 segundos tem-se lambda = 90

Seja $X'$ a v. a. que representa o número de automóveis que entram numa AE num período de 5 minutos (= 300 segundos = 30 × 10 segundos), com $X' \sim P(\lambda' = 10\lambda = 90)$.

Como λ > 20, podemos usar a aproximação à distribuição Normal e

$$X' \overset{\circ}{\sim} N(\mu = \lambda' = 90;\ \sigma = \sqrt{\lambda'} = \sqrt{90}).$$

**Nota**:
Lambda = 90 > 5
Logo pode-se aproximar a distribuição Poisson à distribuição Normal

a) $P(X' > 59) = 1 - P(X' \leq 59) \approx 1 - P\left(Z \leq \frac{59 + 0{,}5 - 90}{\sqrt{90}}\right) = 1 - P(Z \leq -3{,}21) = 1 - \Phi(-3{,}21)$

$= \Phi(3{,}21) = 0{,}9993.$

b) $P(55 \leq X' \leq 80) \approx P\left(\frac{55 - 0{,}5 - 90}{\sqrt{90}} \leq Z \leq \frac{80 + 0{,}5 - 90}{\sqrt{90}}\right) = P(-3{,}74 \leq Z \leq -1{,}00)$

$= \Phi(-1{,}00) - \Phi(-3{,}74) = \big(1 - \Phi(1{,}00)\big) - \big(1 - \Phi(3{,}74)\big) = \Phi(3{,}74) - \Phi(1{,}00)$

$\approx 0{,}9999 - 0{,}8413 = 0{,}1586.$

ProbabilidadesEstatistica2019.pdf

2º semestre – 2018/2019
04/05/2019 – **9:00**

**1.** A rede informática interna de um banco funciona permanentemente. Considere que o tempo (em minuto) de utilização desta rede por um utilizador autenticado é uma variável aleatória $X$ com função de distribuição dada por:

$$F_X(x) = \begin{cases} 1 - e^{-\sqrt{\frac{x}{75}}}, & x > 0 \\ 0, & \text{caso contrário.} \end{cases}$$

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(s)\,ds$$

(a) Obtenha a probabilidade de o tempo de utilização $X$ ser superior a 30 minutos e inferior a 3 horas. (

(b) Tendo em conta que $\overbrace{E(X)}^{\mu} = 150$ e $\overbrace{V(X)}^{\sigma^2} = 112500$, obtenha um valor aproximado para a probabilidade de o tempo médio de $\underbrace{\text{utilização da rede por}}_{X_i \ (i\text{-ésimo})}$ 36 utilizadores autenticados ser superior a 2 horas. (3.0)

# Exercício 1 (a)

$$P(30 < X < 3 \times 60 = 180) = F_X(180) - F_X(30)$$

$$= 1 - e^{-\sqrt{\frac{180}{75}}} - 1 + e^{-\sqrt{\frac{30}{75}}} = e^{-\sqrt{\frac{30}{75}}} - e^{-\sqrt{\frac{180}{75}}}$$

# Exercício 1 (b)

$X_i$ = tempo de utilização de rede do i-ésimo utilizador

$$P\left(\frac{\sum_{i=1}^{36} X_i}{36} > 120\right) = P\left(\sum_{i=1}^{36} X_i > 120 \times 36\right)$$

Porém $X_i$ não está em nenhuma das situações que acabámos de descrever.

Como fazer?

Resposta: Teorema do Limite Central

# Teorema do Limite Central

Teorema: Seja $\{X_i, i=1,\dots n\}$ um conjunto de v.a. independentes e com a mesma distribuição, que não pode ser a distribuição normal. Então

$$\boxed{P\left(\sum_{i=1}^{n} X_i \le x\right) \xrightarrow[n\to\infty]{} \Phi\left(\frac{x-n\mu}{\sigma\sqrt{n}}\right)}$$

onde $\mu = E(X_i)$ e $\sigma^2 = Var(X_i)$, $\forall i$

$$\Leftrightarrow \quad \frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \overset{a}{\sim} \mathcal{N}(0,1)$$

# Exercício 1 (b)

voltando ao exercício:

- $X_i$ = tempo de utilização da rede pelo i-ésimo utilizador
- $X_i \sim ?(\ )$ ; $\underbrace{E(X_i)}_{=\mu} = 150$ ; $\underbrace{var(X_i)}_{=\sigma^2} = 112500$
- $P\left(\sum_{i=1}^{36} X_i > 120 \times 36\right) =$

$$= 1 - P\left(\sum_{i=1}^{36} X_i \leq 120 \times 36\right) = (*)$$

- $n = 36$ [Para efeitos práticos, podemos utilizar o Teorema do Limite Central (TLC) para $n \geq 30$]

$$(*) \simeq 1 - \Phi\left(\frac{120 \times 36 - \overbrace{36 \times 150}^{n\mu}}{\underbrace{\sqrt{112500}\sqrt{36}}_{\sigma\sqrt{n}}}\right)$$

# Exercício 1 (b)

NOTA: se $X_i \sim N(\mu, \sigma^2)$ então não se pode utilizar o Teorema do Limite Central porque em vez de termos:

$$\boxed{P\left(\sum_{i=1}^{n} X_i \leq x\right) \underset{n \to \infty}{\longrightarrow} \Phi\left(\frac{x - n\mu}{\sigma\sqrt{n}}\right)}$$

Vamos ter

$$\boxed{P\left(\sum_{i=1}^{n} X_i \leq x\right) = \Phi\left(\frac{x - n\mu}{\sigma\sqrt{n}}\right)}$$

Slides Professora Claúdia Nunes

**5.20** O tempo (em horas) que João Pestana dorme por noite é uma variável aleatória com distribuição uniforme no intervalo (7,12).

(a) Calcule a probabilidade de João Pestana dormir mais de 11 horas numa noite.

(b) Calcule a probabilidade de, em 20 noites, João Pestana dormir mais de 11 horas em pelo menos 3 dessas noites.

(c) Qual a probabilidade de João Pestana dormir mais de 1100 horas em 100 noites?

# Exercício 5.20 (a)

- **V.a. de interesse**

  $X$ = tempo (em horas) que JP dorme por noite

- **Distribuição de $X$**

$$X \sim \text{Uniforme}(a,b)$$

com

$$\begin{aligned} a &= 7 \\ b &= 12 \end{aligned}$$

- **F.d.p. de $X$**

$$f_X(x) \stackrel{form}{=} \begin{cases} \frac{1}{12-7} = 0.2, & x \in (7,12) \\ 0, & \text{caso contrário} \end{cases}$$

- **Prob. pedida**

$$\begin{aligned} P(X > 11) &= \int_{11}^{+\infty} f_X(x)\,dx \\ &= \int_{11}^{12} 0.2\,dx + \int_{12}^{+\infty} 0\,dx \\ &= (0.2x)|_{11}^{12} \\ &= 0.2. \end{aligned}$$

Slides Professora Sofia Naique

# Exercício 5.20 (a)

X = Tempo (em horas) q dorme numa noite

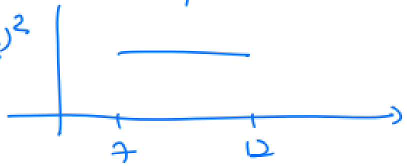
Unif (7, 12)

$f_X(x) = \frac{1}{b-a}, a \leq x \leq b$

$E(X) = \frac{b+a}{2}$ $\quad Var(X) = \frac{(b-a)^2}{12}$

$$f_X(x) = \begin{cases} \frac{1}{5} & x \in (7, 12) \\ 0 & c.c. \end{cases}$$

$E(x) = \frac{7+12}{2} = 9.5$ $\quad var(x) = \frac{(12-7)^2}{12} = \frac{25}{12}$




a) $P(X>11) = \int_{11}^{12} \frac{1}{5}\, dx = \frac{1}{5}$

Slides Professora Claúdia Nunes

# Exercício 5.20 (b)

- **V.a. de interesse**

  $Y$ = no. de noites em que JP dorme mais de 11 horas, num total de 20 noites

- **Distribuição de $Y$**

$$Y \sim \text{binomial}(n,p)$$

com

$$\begin{aligned} n &= 20 \\ p &= P(X > 11) \\ &\stackrel{(a)}{=} 0.2 \end{aligned}$$

- **F.p. de $Y$**

$$P(Y=y) = \binom{20}{y} 0.2^y (1-0.2)^{20-y}, \quad y = 0,1,2,\ldots,20$$

- **Prob. pedida**

$$\begin{aligned} P(Y \geq 3) &= 1 - P(Y < 3) \\ &= 1 - P(Y \leq 2) \\ &= 1 - F_{binomial(20,0.2)}(2) \\ &\underset{tabela/calc}{=} 1 - 0.2061 \\ &= 0.7939 \end{aligned}$$

Slides Professora Sofia Naique

# Exercício 5.20 (c)

- **V.a.**

  $X_i$ = tempo (em horas) que JP dorme na noite $i$, $\quad i = 1,\dots,n$

  $n = 100$

- **Distribuição, valor esperado e variância comuns**

  $X_i \overset{i.i.d.}{\sim} X, \quad i = 1,\dots,n$

  $E(X_i) = E(X) = \mu \overset{form}{=} \frac{a+b}{2} = \frac{7+12}{2} = 9.5, \quad i = 1,\dots,n$

  $V(X_i) = V(X) = \sigma^2 \overset{form}{=} \frac{(b-a)^2}{12} = \frac{(12-7)^2}{12} = \frac{25}{12}, \quad i = 1,\dots,n$

- **V.a. de interesse**

  $S_n = \sum_{i=1}^n X_i$ = tempo (em horas) que JP dorme em $n$ noites

- **Valor esperado e variância de $S_n$**

  $E(S_n) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) \overset{X_i \sim X}{=} \times n\, E(X) = n\, E(X) = n\,\mu$

  $V(S_n) = V\left(\sum_{i=1}^n X_i\right) \overset{X_i\, indep.}{=} \sum_{i=1}^n V(X_i) \overset{X_i \sim X}{=} \times n V(X) = n\, V(X) = n\,\sigma^2$

- **Distribuição aproximada de $S_n$**

  Pelo teorema do limite central (TLC) pode escrever-se

  $$\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} = \frac{S_n - n\,\mu}{\sqrt{n\,\sigma^2}} \overset{a}{\sim} \text{Normal}(0,1).$$

- **Valor aproximado da probabilidade pedida**

  Atendendo a que $n\,\mu = 100 \times 9.5 = 950$ e $n\,\sigma^2 = 100 \times \frac{25}{12} = \frac{625}{3}$, temos

  $$\begin{aligned}
  P(S_n > 1100) &= 1 - P(S_n \le 1100) \\
  &= 1 - P\left(\frac{S_n - n\,\mu}{\sqrt{n\,\sigma^2}} \le \frac{1100 - n\,\mu}{\sqrt{n\,\sigma^2}}\right) \\
  &\overset{TLC}{\simeq} 1 - \Phi\left(\frac{1100-950}{\sqrt{\frac{625}{3}}}\right) \\
  &\simeq 1 - \Phi(10.39) \\
  &< 1 - \Phi(4.09) \\
  &\overset{tabela/calc.}{\simeq} 1 - 0.999978 \\
  &= 0.000022.
  \end{aligned}$$

- [**Obs.**

  Tirando partido do facto da f.d. $\Phi(z)$ ser estritamente crescente e do "último" valor da tabela da f.d. da normal padrão, podemos adiantar que:

  - $10.39 > 4.09 \quad \Rightarrow \quad \Phi(10.39) > \Phi(4.09) \overset{tabela/calc.}{\simeq} 0.999978;$
  - $-10.39 < -4.09 \quad \Rightarrow \quad \Phi(-10.39) < \Phi(-4.09) = 1 - \Phi(4.09) \overset{tabela/calc.}{\simeq} 1 - 0.999978 = 0.000022.$]

Slides Professora Sofia Naique

# Obrigada!

## Questões?