



Lisbon School  
of Economics  
& Management  
Universidade de Lisboa



# Estatística I

Licenciatura em Gestão e Licenciatura em Finanças  
2.º Ano/2.º Semestre  
2023/2024

# Aulas Teóricas N.ºs 20 e 21 (Semana 12)

**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)

- **Capítulo 2:**  
Probabilidades

## Aulas Teóricas (Semanas 3 a 5)

- **Capítulo 3:** Variáveis  
Aleatórias  
Unidimensionais

## Aulas Teóricas (Semanas 5 a 7)

- **Capítulo 4:** Variáveis  
Aleatórias  
Multidimensionais

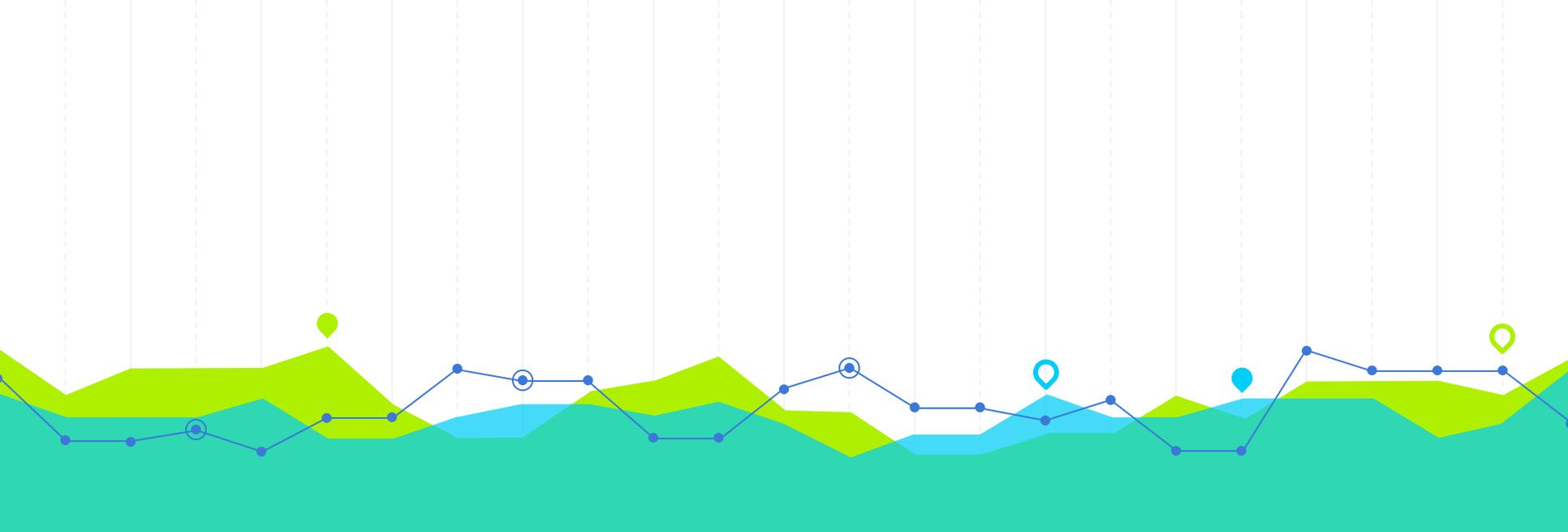
## Aulas Teóricas (Semanas 8 a 13)

- **Capítulo 5:**  
Distribuições Teóricas
- **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>



# Teorema do Limite Central

1

# 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  $X_1, \dots, X_n$  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  $n \rightarrow +\infty$

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

ou seja,

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

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

# TLC

## Teorema do Limite Central

De modo equivalente, se tem

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

ou seja,

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

Slides da Professora Conceição Amado

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

Observar que:

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

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

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

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

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

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

Formulário

AMOSTRAGEM. DISTRIBUIÇÕES POR AMOSTRAGEM

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

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

# TLC: Aproximações

## Observações:

- O TLC deve ser usado **apenas** quando as v.a.'s  $X_1, \dots, X_n$  **não têm distribuição normal!**
- A demonstração do teorema exige algumas ferramentas matemáticas avançadas.
- As v.a.  $X_1, \dots, X_n$  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,  $X_1, X_2, \dots, X_n, \dots$ , com distribuição de Bernoulli de média  $E(X_i) = \theta$  e, portanto,  $\text{Var}(X_i) = \theta(1 - \theta)$ , tem-se,

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

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

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

# Aproximações Baseadas no TLC: Poisson - Normal

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

A aproximação será tanto melhor quanto maior for  $\lambda$ .

# 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  $0, \Delta, 2\Delta, 3\Delta, \dots$ ), com  $E(X) = \mu_X$  e  $Var(X) = \sigma_X^2$ . Então a correção de continuidade é efetuada da seguinte forma:

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

# 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

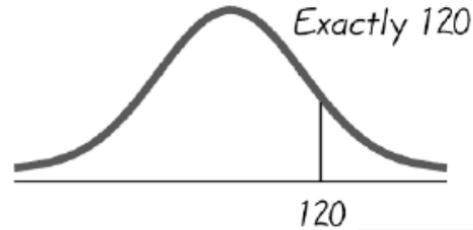
$$P(X = k) \approx P(k - 0,5 \leq X < k + 0,5).$$

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

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

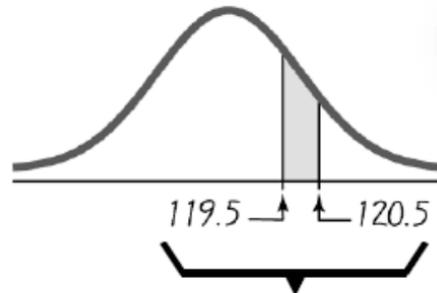
# Aproximações Baseadas no TLC: Correção de Continuidade

$x =$ exatamente 120



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

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



Intervalo que representa o valor discreto 120

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

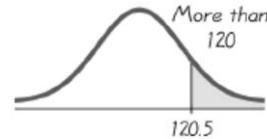
# Aproximações Baseadas no TLC: Correção de Continuidade

$X = \underline{\text{pelo menos}} 120$   
 $= 120, 121, 122, \dots$



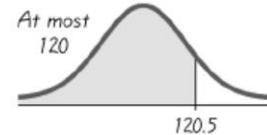
$$P(X \geq 120) \sim P(\tilde{X} \geq 120-0,5)$$

$X = \underline{\text{mais do que}} 120$   
 $= 121, 122, 123, \dots$



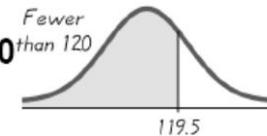
$$P(X > 120) \sim P(\tilde{X} > 120+0,5)$$

$X = \underline{\text{no máximo}} 120$   
 $= 0, 1, \dots 118, 119, 120$



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

$X = \underline{\text{menos do que}} 120$   
 $= 0, 1, \dots 118, 119$



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

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

# 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:

- $P(X = a) \approx P(a - \delta < \tilde{X} < a + \delta)$
- $P(a < X < b) \approx P(a + \delta \leq \tilde{X} \leq b - \delta)$
- $P(X \leq a) \approx P(\tilde{X} \leq a + \delta)$
- $P(X < a) \approx P(\tilde{X} < a - \delta) = P(\tilde{X} \leq a - \delta)$  (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.

# TLC

## Formulário

### TEOREMA DO LIMITE CENTRAL E COROLÁRIOS

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

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

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

Corolário: Sendo  $X \sim \text{Po}(\lambda)$ , quando  $\lambda \rightarrow +\infty \Rightarrow \frac{X - \lambda}{\sqrt{\lambda}} \stackrel{a}{\sim} N(0,1)$

Correcção de continuidade:  $P(a \leq X \leq b) \approx \Phi\left(\frac{b + \frac{1}{2} - \lambda}{\sqrt{\lambda}}\right) - \Phi\left(\frac{a - \frac{1}{2} - \lambda}{\sqrt{\lambda}}\right)$ , 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  $0.1\text{kg}^2$ . 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 paleta (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:

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

$$P(S > 510) = 1 - P(S \leq 510) = 1 - P\left(\frac{S - 500}{\sqrt{100}} \leq \frac{510 - 500}{\sqrt{100}}\right) \stackrel{T.L.C.}{\approx} \\ \stackrel{T.L.C.}{\approx} 1 - \Phi(1) \approx 0.1587.$$

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:

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

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

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

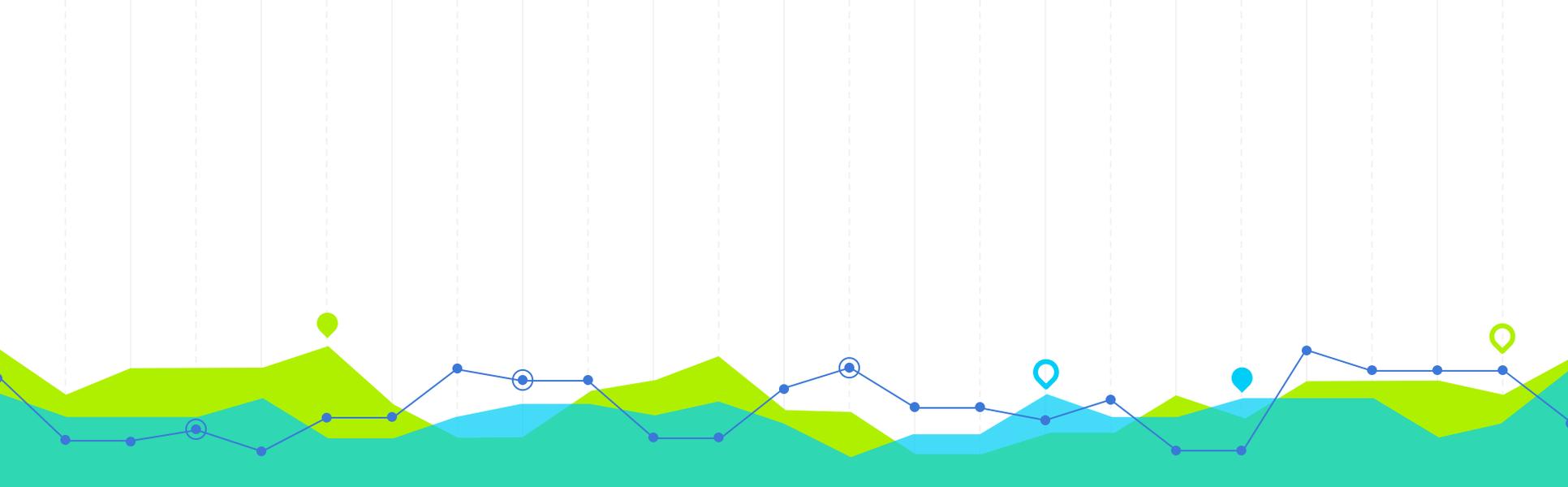
$$= 1 - \Phi(2.59) \approx 0.0048$$

### Nota:

Lambda = 1500 > 5

Logo pode-se aproximar a distribuição Poisson à distribuição Normal

Slides da Professora Conceição Amado



# Teorema do Limite Central: Exercícios

# 2

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,

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

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

Seja  $\bar{X} = \frac{1}{30}(X_1 + \dots + X_{30})$  a v.a. que representa a média dos 30 tempos de viagem de autocarro entre Lisboa e Évora.

Pelo corolário do T. L. C.,

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

Logo,

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

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  $X_i \sim \text{Exp}\left(\lambda = \frac{1}{15}\right)$ . Logo,

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

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

Seja  $S_{100} = X_1 + \dots + X_{100}$  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.,

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

Logo,

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

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

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

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

$$\begin{aligned} \text{b) } P(16 < X < 45) &\approx P\left(\frac{16 + 0,5 - 20}{4} < Z < \frac{45 - 0,5 - 20}{4}\right) = P(-0,88 < Z < 6,13) \\ &= \Phi(6,13) - \Phi(-0,88) = \Phi(6,13) - (1 - \Phi(0,88)) \approx 1 - (1 - 0,8106) = 0,8106. \end{aligned}$$

### Nota:

$$n \times p = 20 > 5$$

$$n \times p \times (1-p) = 16 > 5$$

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 =  $30 \times 10$  segundos), com  $X' \sim P(\lambda' = 10\lambda = 90)$ .

Como  $\lambda > 20$ , podemos usar a aproximação à distribuição Normal e

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

$$\begin{aligned} \text{a) } P(X' > 59) &= 1 - P(X' \leq 59) \approx 1 - P\left(Z \leq \frac{59 + 0,5 - 90}{\sqrt{90}}\right) = 1 - P(Z \leq -3,21) = 1 - \Phi(-3,21) \\ &= \Phi(3,21) = 0,9993. \end{aligned}$$

$$\begin{aligned} \text{b) } P(55 \leq X' \leq 80) &\approx P\left(\frac{55 - 0,5 - 90}{\sqrt{90}} \leq Z \leq \frac{80 + 0,5 - 90}{\sqrt{90}}\right) = P(-3,74 \leq Z \leq -1,00) \\ &= \Phi(-1,00) - \Phi(-3,74) = (1 - \Phi(1,00)) - (1 - \Phi(3,74)) = \Phi(3,74) - \Phi(1,00) \\ &\approx 0,9999 - 0,8413 = 0,1586. \end{aligned}$$

**Nota:**

$\lambda = 90 > 5$

Logo pode-se aproximar a distribuição Poisson à distribuição Normal

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:

$$F_X(x) = \begin{cases} 1 - e^{-\sqrt{\frac{x}{30}}}, & x > 0 \\ 0, & \text{caso contrário.} \end{cases}$$

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

- (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  $E(X) = 150$  e  $V(X) = 112500$ , obtenha um valor aproximado para a (3.0) probabilidade de o tempo médio de utilização da rede por 36 utilizadores autenticados ser superior a 2 horas.



## Exercício 1 (a)

$$\begin{aligned} P(30 < X < 3 \times 60 = 180) &= F_X(180) - F_X(30) \\ &= 1 - e^{-\sqrt{\frac{180}{75}}} - 1 + e^{-\sqrt{\frac{30}{75}}} = e^{-\sqrt{\frac{30}{75}}} - e^{-\sqrt{\frac{180}{75}}} \end{aligned}$$

## 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.c. independentes e com a mesma distribuição, que não pode ser a distribuição é normal. Então

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

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

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

## Exercício 1 (b)

Voltando ao exercício:

- $X_i$  = tempo de utilização de rede pelo  $i$ -ésimo utilizador
- $X_i \sim \text{CN}(\mu)$ ;  $E(X_i) = 150$ ;  $\text{var}(X_i) = 112500 = \sigma^2$
- $P\left(\sum_{i=1}^{36} X_i > 120 \times 36\right) =$   
 $= 1 - P\left(\sum_{i=1}^{36} X_i \leq 120 \times 36\right) = \textcircled{*}$
- $n = 36$  [Para efeitos práticos, podemos utilizar o Teorema do Limite Central para  $n \geq 30$ ] (TLC)

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

## Exercício 1 (b)

NOTA : se  $X_i \sim \text{CN}(\mu, \sigma^2)$  então não se pode utilizar o Teorema do Limite Central porje em vez de termos:

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

Vamos ter

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

**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

$$a = 7$$

$$b = 12$$

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

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

- **Prob. pedida**

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

## Exercício 5.20 (a)

$X =$  Tempo (em horas)  $\bar{y}$  de uma viagem noturna  
uniforme  $(7, 12)$

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

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

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

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



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

# 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

$$n = 20$$

$$p = P(X > 11)$$

$$\stackrel{(a)}{=} 0.2$$

- **Ep. de  $Y$**

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

- **Prob. pedida**

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

## Exercício 5.20 (c)

- **V.a.**

$X_i$  = tempo (em horas) que JP dorme na noite  $i$ ,  $i = 1, \dots, n$   
 $n = 100$

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

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

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

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

- **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  $S_n$**

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

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

- **Distribuição aproximada de  $S_n$**

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

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

- **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

$$\begin{aligned} P(S_n > 1100) &= 1 - P(S_n \leq 1100) \\ &= 1 - P\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \leq \frac{1100 - n\mu}{\sqrt{n\sigma^2}}\right) \\ &\stackrel{\text{TLC}}{\approx} 1 - \Phi\left(\frac{1100 - 950}{\sqrt{\frac{625}{3}}}\right) \\ &\approx 1 - \Phi(10.39) \\ &< 1 - \Phi(4.09) \\ &\stackrel{\text{tabela/calcul.}}{\approx} 1 - 0.999978 \\ &= 0.000022. \end{aligned}$$

- **[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:

- $10.39 > 4.09 \Rightarrow \Phi(10.39) > \Phi(4.09) \stackrel{\text{tabela/calcul.}}{\approx} 0.999978;$
- $-10.39 < -4.09 \Rightarrow \Phi(-10.39) < \Phi(-4.09) = 1 - \Phi(4.09) \stackrel{\text{tabela/calcul.}}{\approx} 1 - 0.999978 = 0.000022.]$



# Inferência Estatística

# 3

Conceitos: Universo, Amostra Aleatória, Distribuição de uma Amostra Aleatória e Estatísticas

# Probabilidade vs. Inferência Estatística

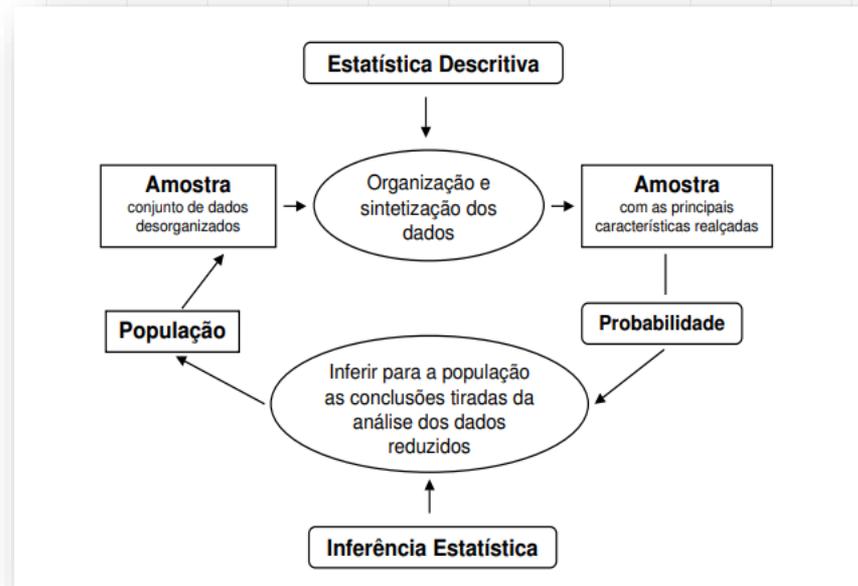
Pode dizer-se que a **Probabilidade e a Inferência** têm objectivos diferentes: enquanto na Probabilidade se parte de um dado esquema ou modelo para calcular a probabilidade de certos resultados ou acontecimentos se realizarem; na Inferência parte-se de dados ou observações e procura saber-se ou inferir-se algo sobre o modelo.

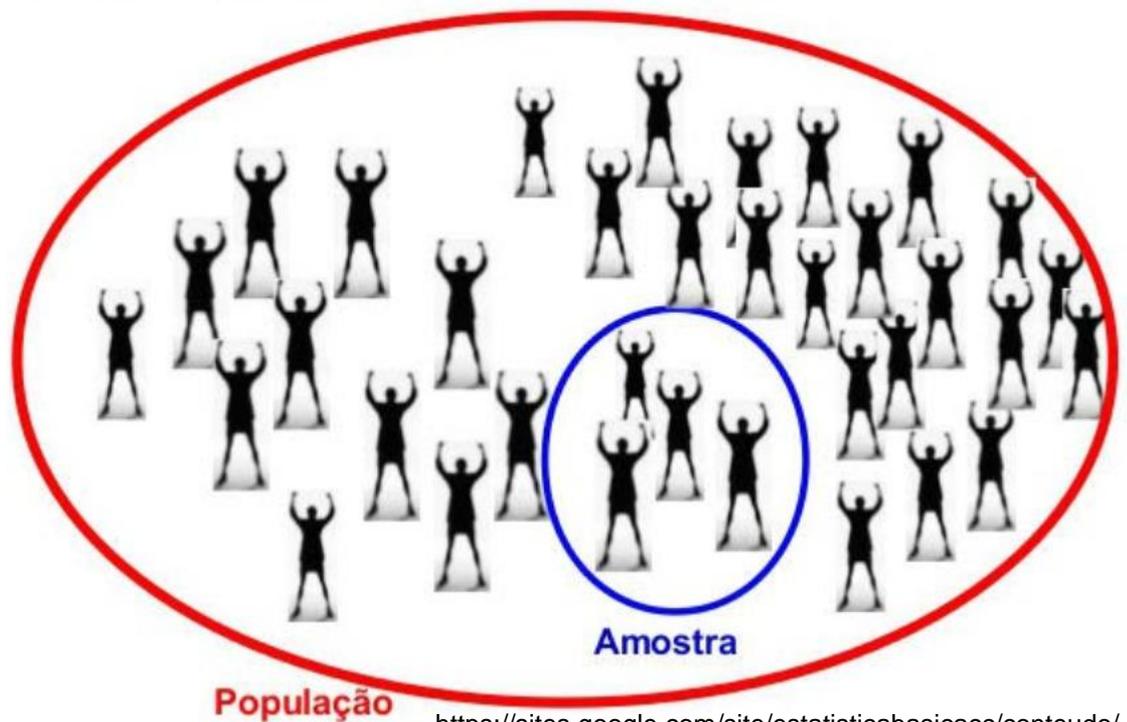
A Inferência é a “passagem do particular ao geral.”

A Inferência Estatística tem como objectivo **definir procedimentos** que aplicados a uma amostra extraída da população, **nos permitam estimar parâmetros desconhecidos dessa população ou algo sobre o modelo da população.**

De facto uma amostra particular é apenas uma das muitas amostras (em  $n^\circ$  infinito se a população for infinita) que se podem obter por um **processo de amostragem.**

# Estatística Descritiva vs. Inferência Estatística





<https://sites.google.com/site/estatisticabasicacc/conteudo/>

➤ **POPULAÇÃO:**

- é uma coleção completa de todos os elementos a serem estudados

➤ **AMOSTRA:**

- é um subconjunto da população

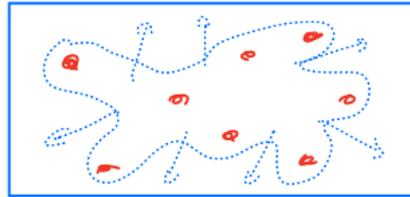
➤ **CENSO:**

- é uma coleção de dados relativos a todos os elementos de uma população:

<https://slideplayer.com.br/slide/2627398>

# Amostra Aleatória

Perante "falta de informação" sobre a v.c. (população),  
vamos recolher uma amostra, estudar intensamente  
esta amostra e tentar extrapolar p/ a população.



$\Omega$

• - elementos da  
população que  
vamos estudar  
↓  
amostra  
( $x_1, x_2, \dots, x_n$ )

# População vs Amostra Aleatória

população

(X)

v.c. interesse

→ recolher um subconjunto →  
representativo de população

①

$(X_1, X_2, \dots, X_n)$

(amostra aleatória)

②

estudar a amostra,  
através de estatística  
descritiva

③

inferir ("extrapolar")  
os resultados de amostra  
p/ a população.

# Amostragem

① Amostragem (Regras p/ assegurar a "bondade" dos resultados)

Amostragem é representativa de população se tiver as seguintes características:

$(X_1, X_2, \dots, X_n)$  : conjunto de variáveis aleatórias independentes entre si e têm todas a mesma distribuição, que é a distribuição da v.c. (população)  
X

Slides Professora Cláudia Nunes

a.a. = amostragem aleatória  
i.i.d. = independentes e idêntico/distribuídas

# Distribuição de uma a.a.: V.a.'s Discretas

Qual é a distribuição da a.a.  $(X_1, \dots, X_n)$ , sendo as v.a.'s  $X_i$ 's ( $i = 1, \dots, n$ ) iid a uma v.a. discreta  $X$ ?



# Distribuição de uma a.a.: V.a.'s Discretas

1ª questão: Qual é distribuição de uma c. c.?

i) se  $X$  for uma v. c. discreta

$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n)$  [função de prob.  
conjunta de a.a.]

$$\stackrel{\substack{= \\ \downarrow \\ \text{independentes}}}{=} P(X_1 = x_1) P(X_2 = x_2) \dots P(X_n = x_n)$$

$$\stackrel{\substack{= \\ \downarrow \\ \text{ident. distribuídos}}}{=} P(X = x_1) P(X = x_2) \dots P(X = x_n)$$

$$= \prod_{i=1}^n P(X = x_i)$$

## Distribuição de uma a.a.: V.a.'s Poisson

Determine a seguinte função de probabilidade conjunta da a.a.  $(X_1, X_2, X_3)$ :  $P(X_1=4, X_2=3, X_3=5)$ , sendo  $X_i$ 's ( $i = 1, 2, 3$ ) iid a  $X \cap P(\lambda)$ .



# Distribuição de uma a.a.: V.a.'s Poisson

$$\begin{aligned} \text{exemplo: } X &\sim \text{Poisson}(\lambda) \quad \lambda = ? \\ P(X_1 = 4, X_2 = 3, X_3 = 5) &= \prod_{i=1}^3 \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} \\ &= \frac{e^{-\lambda} \lambda^4}{4!} \frac{e^{-\lambda} \lambda^3}{3!} \frac{e^{-\lambda} \lambda^5}{5!} = \frac{e^{-3\lambda} \lambda^{12}}{4! 3! 5!} \end{aligned}$$

Slides Professora Cláudia Nunes

# Distribuição de uma a.a.: V.a.'s Contínuas

Qual é a distribuição da a.a.  $(X_1, \dots, X_n)$ , sendo  $X_i$ 's ( $i = 1, \dots, n$ ) iid a uma v.a. Contínua  $X$ ?



# Distribuição de uma a.a.: V.a.'s Contínuas

ii) se  $X$  for uma v.c. contínua

$f_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n)$  [função de densidade de probabilidade conjunta]

$\underset{\substack{= \\ \text{independentes}}}{=} f_{X_1}(x_1) f_{X_2}(x_2) \dots f_{X_n}(x_n)$

$= f_X(x_1) f_X(x_2) \dots f_X(x_n)$

idênticas e distribuídas

$= \prod_{i=1}^n f_X(x_i)$

## Distribuição de uma a.a.: V.a.'s Exponenciais

Determine a seguinte função densidade de probabilidade conjunta da a.a.  $(X_1, X_2, X_3, X_4)$ :  $f(1.3, 0.5, 2.5, 0.78)$ , sendo  $X_i$ 's ( $i = 1, 2, 3, 4$ ) iid a  $X \cap \text{Exp}(\lambda)$ .



# Distribuição de uma a.a.: V.a.'s Exponenciais

Exemplo:  $X \sim \text{Exp}(\mu)$

$$X_1 = 1.3 ; X_2 = 0.5 ; X_3 = 2.5 ; X_4 = 0.78$$

$$f_{X_1, X_2, X_3, X_4}(1.3, 0.5, 2.5, 0.78) = \mu e^{-\mu \times 1.3} \times \mu e^{-\mu \times 0.5}$$

$$\times \mu e^{-\mu \times 2.5} \times \mu e^{-\mu \times 0.78}$$

$$= \mu^4 e^{-\mu(1.3 + 0.5 + 2.5 + 0.78)} //$$

# Inferência Estatística

## ③ Inferência estatística

- $X \sim \mathcal{D}$  (v.c. de interesse)
- Com distribuição conhecida  $\mathcal{D}$
- Todos ou alguns parâmetros de distribuição desconhecidos

parâmetros desconhecidos ( $\theta$ )

Objetivo: "adivinhar" o valor de  $\theta$ , com base na amostra  
estimar  $\theta$ , com base na amostra

↓  
ESTIMACÃO { Pontual (cap. 6)  
Intervalar (cap. 7)

Slides Professora Cláudia Nunes

# Estatística

i) ESTATÍSTICA: uma estatística é uma função (qualquer) que depende da amostra e não depende de nenhum parâmetro desconhecido.

Exemplo:  $X \sim \text{Poisson}(\lambda)$        $\lambda = ?$

a.a.  $(X_1, X_2, \dots, X_{10})$

$T_1 = \bar{X} = \frac{1}{10} \sum_{i=1}^{10} X_i$  : É UMA ESTATÍSTICA

$T_2 = 3$  : É UMA ESTATÍSTICA

$T_3 = \frac{1}{10} \sum_{i=1}^{10} (X_i - \lambda)^2$  NÃO É ESTATÍSTICA!!

# Estatística

## DEF. 4: ESTATÍSTICA:

Uma estatística é uma função das variáveis observáveis, e é por si própria uma variável observável que não contém qualquer parâmetro desconhecido.

**EXEMPLO 5:** Seja a a.a.  $[X_1, X_2, \dots, X_n]$ . A média amostral  $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$  é uma estatística.

**EXEMPLO 6:** Seja a v.a.  $X \sim N(\mu, \sigma^2)$  com  $\mu$  e  $\sigma^2$  desconhecidos, então  $x - \mu$  não é estatística porque não é observável, é função do parâmetro desconhecido  $\mu$ .

**EXEMPLO 6:** Seja uma v.a. com distribuição  $N(\mu; \sigma^2)$  Quais são Estatísticas?

a)  $X^2 - \mu$

d)  $X - 4$

b)  $\frac{X}{\sigma^2}$

e)  $X - \log X^3$

c)  $X^2 - 3$

Quais destas funções são estatísticas?

# Média e Variância Amostrais

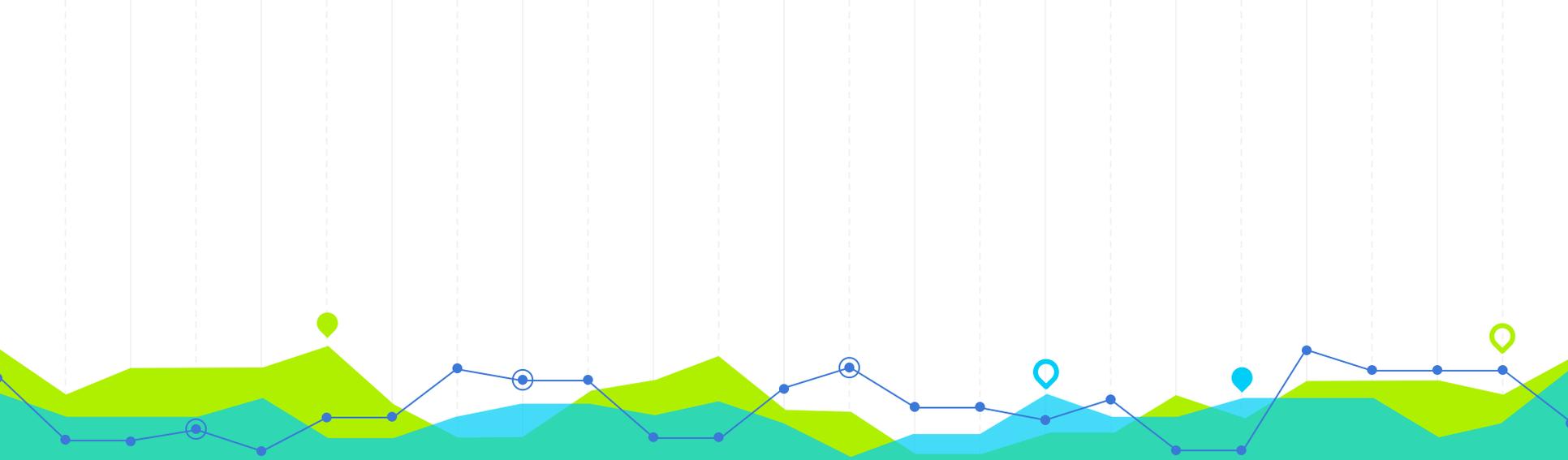
② Das características de amostra frequentemente calculadas são:

média amostral:  $\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i$

variância amostral:  $s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$   
 $= \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$  \*

⊙ porque:  $\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 =$   
 $= \frac{1}{n} \sum_{i=1}^n (x_i^2 + \bar{x}^2 - 2x_i\bar{x}) = \frac{1}{n} \left[ \sum_{i=1}^n x_i^2 + n\bar{x}^2 - 2\bar{x} \sum_{i=1}^n x_i \right]$   
 $= \frac{1}{n} \left[ \sum_{i=1}^n x_i^2 + n\bar{x}^2 - 2n\bar{x}^2 \right]$   
 $= \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2$

variância amostral corrigida:  $s_{n-1}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$



# Média Amostral

Distribuições de Amostragem

# 4

# Média Amostral

Seja  $X_1, X_2, \dots, X_n$ , uma a. a. dum dada população com média  $\mu$  e variância  $\sigma^2$ . A **média amostral** é definida por:

$$\bar{X} = \sum_{i=1}^n \frac{X_i}{n}.$$

## Propriedades:

- $\mu_{\bar{X}} = E(\bar{X}) = \mu;$
- $\sigma_{\bar{X}}^2 = Var(\bar{X}) = \frac{1}{n}\sigma^2.$

# Média Amostral: Variância Conhecida

Se a distribuição da população é Normal com desvio padrão  $\sigma$  conhecido, então  $\bar{X} \sim N\left(\mu; \frac{\sigma}{\sqrt{n}}\right)$ , ou seja,

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0; 1).$$

Se a distribuição da população não for Normal, mas a amostra é de grande dimensão então, pelo corolário do T. L. C., vem

$$Z \overset{\sim}{\sim} N(0; 1).$$

# Média Amostral: Variância Desconhecida

Se a população é Normal mas o desvio padrão  $\sigma$  é desconhecido, e não rejeitando a hipótese de independência das distribuições por amostragem da média e da variância da amostra, então tem-se que:

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}.$$

Se a distribuição da população não for Normal, mas a amostra for de grande dimensão então, por extensão do corolário do T. L. C.,

$$\bar{X} \overset{\circ}{\sim} N\left(\mu; \frac{S}{\sqrt{n}}\right), \text{ ou seja, } Z = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \overset{\circ}{\sim} N(0; 1).$$

*Observação:* Para valores elevados de  $n$ , a distribuição  $t$ -Student toma valores muito próximos aos da  $N(0; 1)$ . Existem alguns programas estatísticos (por exemplo, SPSS) que, nestas condições, não utilizam a distribuição Normal mas a  $t$ -Student.

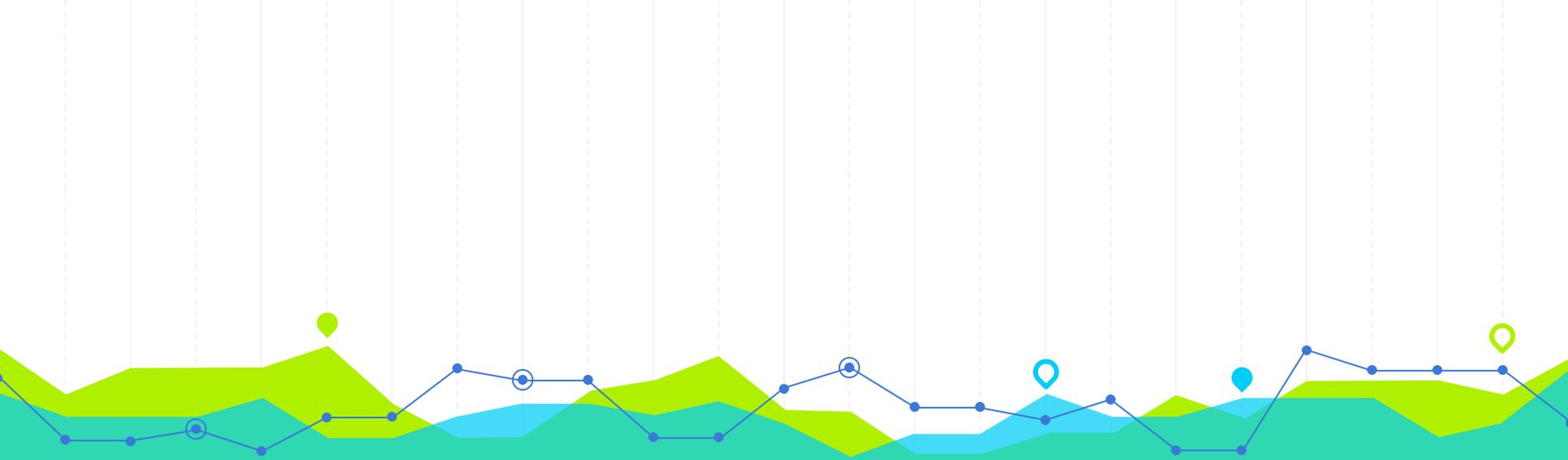
# Distribuições de Amostragem

## Formulário

### AMOSTRAGEM. DISTRIBUIÇÕES POR AMOSTRAGEM

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

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



# Média Amostral: Exercícios

Distribuições de Amostragem

# 5

Um fabricante de automóveis defende que o novo modelo que vai ser lançado no próximo mês gasta em média 9,7 litros aos 100 km, em circuito urbano, com desvio padrão de 1 litro. Admita que o consumo segue uma distribuição Normal.

- a) Qual a probabilidade de numa amostra aleatória de 20 automóveis o gasto médio ser superior a 10 litros.
- b) Qual a deverá ser a dimensão da amostra para obter, com pelo menos 90% probabilidade, um gasto médio inferior a 10 litros.

[ProbabilidadesEstatistica2019.pdf](#)



# Exercício: Média Amostral com Variância Conhecida

Seja  $X$  a v.a. que representa o número de litros consumidos pelo automóvel aos 100 km, em circuito urbano, com  $X \sim N(\mu = 9,7; \sigma = 1)$ .

a)  $n = 20$ .

$$\begin{array}{l} X \text{ dist. Normal} \\ \sigma \text{ conhecido} \end{array} \left| \Rightarrow Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0; 1). \right.$$

$$\begin{aligned} P(\bar{X} > 10) &= 1 - P(\bar{X} \leq 10) = 1 - P\left(Z \leq \frac{10 - 9,7}{\frac{1}{\sqrt{20}}}\right) = 1 - P(Z \leq 1,34) = 1 - \Phi(1,34) \\ &= 1 - 0,9099 = 0,0901. \end{aligned}$$

b)  $n = ?$

$$P(\bar{X} < 10) \geq 0,9 \Leftrightarrow P\left(Z < \frac{10 - 9,7}{\frac{1}{\sqrt{n}}}\right) \geq 0,9 \Leftrightarrow \Phi(0,3\sqrt{n}) \geq 0,9$$

$$\text{como } \Phi(1,282) = 0,9$$

$$\Leftrightarrow 0,3\sqrt{n} \geq 1,282 \Leftrightarrow \sqrt{n} \geq 4,2733 \Rightarrow n \geq 4,2733^2 = 18,3 \Rightarrow n \geq 19.$$

Um fabricante de automóveis defende que o novo modelo que vai ser lançado no próximo mês gasta em média 9,7 litros aos 100 km, em circuito urbano, e o desvio padrão é desconhecido. Através de um esquema de amostragem estimou-se tal desvio padrão como sendo  $s = 1$  litro. Admitindo que o consumo segue uma distribuição Normal, qual a probabilidade de, numa amostra aleatória de 20 automóveis, o consumo médio amostral ser superior a 10 litros? E inferior a 8,9 litros?

[ProbabilidadesEstadistica2019.pdf](#)



# Exercício: Média Amostral com Variância Desconhecida

Seja  $X$  a v.a. que representa o número de litros consumidos pelo automóvel aos 100 km, em circuito urbano, com  $X \sim N(\mu = 9,7; \sigma = ?)$ .

$n = 20$ .

$$\begin{array}{l} X \text{ dist. Normal} \\ \sigma \text{ desconhecido} \end{array} \left| \Rightarrow T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1=19}.$$

$$P(\bar{X} > 10) = 1 - P(\bar{X} \leq 10) = 1 - P\left(T \leq \frac{10 - 9,7}{\frac{1}{\sqrt{20}}}\right) = 1 - P(T \leq 1,342) \approx 1 - 0,9 = 0,1.$$

$$P(\bar{X} < 8,9) = P\left(T < \frac{8,9 - 9,7}{\frac{1}{\sqrt{20}}}\right) = P(T < -3,578) = 1 - P(T < 3,578) \approx 1 - 0,999 = 0,001.$$

# Obrigada!

Questões?

