



Lisbon School  
of Economics  
& Management  
Universidade de Lisboa

# Estatística II

Licenciatura em Gestão do Desporto  
2.º Ano/2.º Semestre  
2023/2024

# Aulas Teórico-Práticas N.ºs 21 e 22 (Semana 13)

**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 Teórico-Práticas (Semanas 1 a 5)

- **Capítulo 1:** Revisões e Distribuições de Amostragem

## Aulas Teórico-Práticas (Semanas 5 a 7)

- **Capítulo 2:** Estimação

## Aulas Teórico-Práticas (Semanas 7 a 9)

- **Capítulo 3:** Testes de Hipóteses

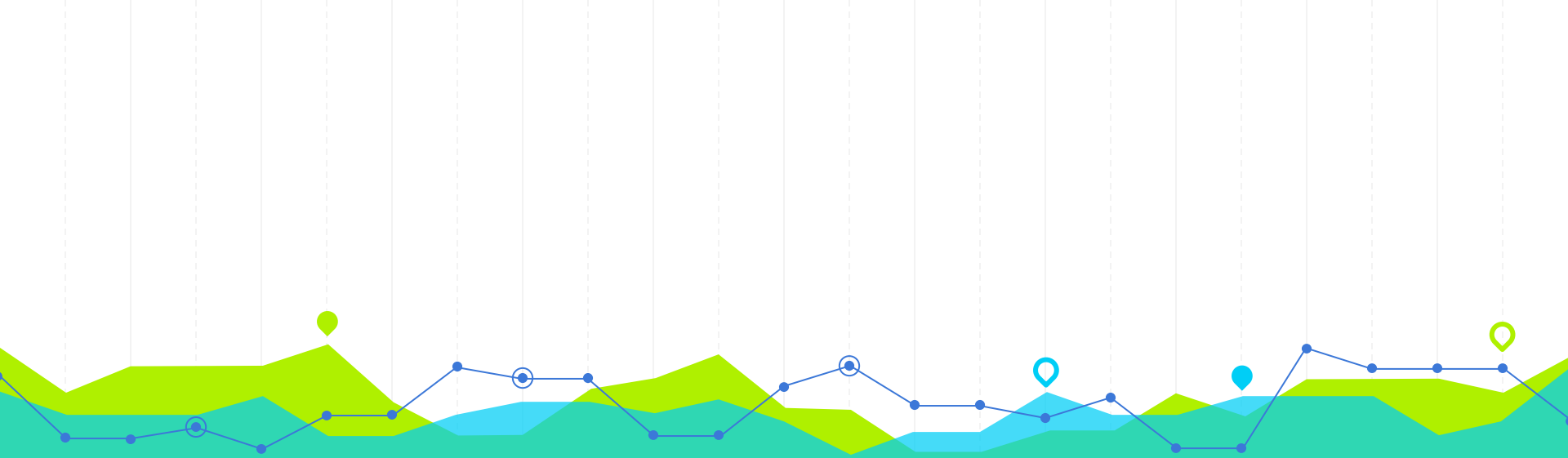
## Aulas Teórico-Práticas (Semanas 10 a 13)

- **Capítulo 4:** Modelo de Regressão Linear Múltipla

**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>



# Pré-Teste

1

**1. O número de automóveis que passam em cada período de um minuto num determinado ponto de uma autoestrada é uma variável aleatória com distribuição de Poisson de valor médio  $\lambda$ . Tendo por base a contabilização do número de automóveis que passam nesse ponto da autoestrada em cada um de 100 intervalos de um minuto, selecionados ao acaso:**

- 1.1. Deduza o estimador de máxima verosimilhança (EMV) de  $\lambda$ . [2.0]
- 1.2. Será que o EMV de  $\lambda$  é centrado? [1.0]
- 1.3. Determine o erro quadrático médio do EMV de  $\lambda$ . [1.0]



**2. Um vendedor de uma marca de baterias, usadas em *tablets*, mediu os tempos (em horas) de autonomia de 16 dessas baterias,  $x_i$ ,  $i = 1, 2, \dots, 16$ , tendo obtido  $\sum_{i=1}^{16} x_i = 79.5$  e  $\sum_{i=1}^{16} x_i^2 = 395.65$ . Admitindo que o tempo de autonomia dessas baterias tem distribuição normal:**

- 2.1. Obtenha um intervalo de confiança (IC) a 95%, para o tempo médio de autonomia das baterias em questão. [2.0]
- 2.2. A amplitude de um intervalo de confiança a 95% para o valor médio (desvio padrão populacional desconhecido) de uma variável com distribuição normal não é influenciada por: [1.0]
- A) A dimensão da amostra.
  - B) O nível/grau de confiança.
  - C) O desvio padrão da amostra.
  - D) A média da amostra.
  - E) O nível de significância 0,05.
- 2.3. Teste ao nível de significância de 5% se a variância do tempo de autonomia das baterias em questão é superior a 0.06. [2.0]
- 2.4. Num teste estatístico, sabe-se que: [1.0]
- A) A região de rejeição não depende do nível de significância.
  - B) O valor-p não depende do nível de significância.
  - C) A probabilidade de erro tipo I é igual à potência do teste.
  - D) A probabilidade de erro tipo II é igual à potência do teste.
  - E) Se o valor-p é superior ao nível de significância ( $\alpha$ ), então rejeita-se a hipótese nula para  $\alpha$ .



**3. O registo, efetuado ao longo de um conjunto de várias semanas, de faltas ao trabalho dos funcionários de uma empresa conduziu aos seguintes resultados:**

Dias	Segunda	Terça	Quarta	Quinta	Sexta
Nº de Faltas	23	17	14	20	26

3.1. Teste a hipótese de as faltas ao trabalho dos funcionários da empresa se distribuírem uniformemente pelos 5 dias úteis da semana. Decida com base no valor-p. [2.0]

3.2. Uma das condições de aplicabilidade do teste de hipóteses referido na alínea anterior é: [1.0]

- A) A variância tem que ser homogénea nos grupos de estudo.
- B) Os valores esperados têm que ter uma distribuição normal.
- C) Pelo menos 80% dos valores observados têm que ser superiores a 5.
- D) Pelo menos 80% dos valores esperados têm que ser superiores a 5.
- E) A dimensão da amostra tem que ser superior a 20.



4. Este estudo pretende avaliar o impacto de alguns fatores que explicam a distância percorrida pelas famílias durante o período de férias. O modelo de regressão linear de referência é o seguinte:

$$dist = \beta_1 + \beta_2rend + \beta_3idade + \beta_4criancas + u$$

onde:

- *dist* – distância percorrida nas férias, em quilómetros;
- *rend* – rendimento familiar anual, em milhares de euros;
- *idade* – idade média dos adultos no agregado familiar, em anos;
- *criancas* =  $\begin{cases} 1, & \text{se existem crianças no agregado familiar} \\ 0, & \text{caso contrário} \end{cases}$

O *output* do EXCEL com o modelo estimado é o seguinte:

<i>Regression Statistics</i>	
Multiple R	0.607
R Square	0.369
Adjusted R Square	0.359
Standard Error	681.426
Observations	194





## ANOVA

	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	3	51585572.745	17195190.915	37.031	0.000
Residual	190	88224848.246	464341.307		
Total	193	139810420.991			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>p-value</i>
Intercept	-431.113	260.136	-1.657	0.099
rend	24.724	2.729	9.059	0.000
idade	17.383	5.543	3.136	0.002
crianças	-188.638	115.240	-1.637	0.103

4.1. Interprete os efeitos estimados do rendimento e da existência de crianças sobre a distância percorrida nas férias. [0.5]

4.2. O que pode concluir quanto à significância estatística da variável binária? [1.0]

- A) É significativa para 5%.
- B) É significativa para 10%.
- C) É significativa para 5%, mas não para 10%.
- D) É significativa para 10%, mas não para 5%.
- E) Não é significativa para 10%.



# Obrigada!

Questões?

