



**ESTATÍSTICA II (Eco+Fin) - 2º Ano/2º sem.**  
**Exame Época Recurso 29 Junho 2023**

**2 horas. (200 pontos)**

**Nome:** \_\_\_\_\_ **nº** \_\_\_\_\_

1 a) 10	2 a) 15	3 a) 15	4) 25	5 a) 15
1 b) 15	2 b) 5	3 b) 5		5 b) 10
1 c) 10	2 c) 10			5 c) 15
	2 d) 15			5 d) 15
				5 e) 20
<b>35</b>	<b>45</b>	<b>20</b>	<b>25</b>	<b>75</b>
<b>Total:</b>				

1. O tempo que um aluno leva a responder a uma pergunta do exame é uma variável aleatória com distribuição exponencial de parâmetro  $\lambda$ .  
 Numa amostra casual de 30 observações verificou-se um total de 450 minutos.

a) Obtenha o estimador do método dos momentos para  $\lambda$ . **[10]**

b) Obtenha o estimador para  $\lambda$  pelo método da máxima verosimilhança. **[15]**

- c) Usando a estimativa de  $\lambda$ , calcule a probabilidade de uma pergunta ser resolvida em menos de 15 min. **[10]**

2. Admita que o tempo de reparação de certo tipo de máquinas,  $X$ , segue uma distribuição normal de parâmetros desconhecidos.

A fim de estimar esses parâmetros recolheu-se uma amostra aleatória de tempos de reparação (em minutos). Os dados são os seguintes:

$$n = 10, \sum x_i = 956, \sum x_i^2 = 91607$$

- a) Construa um intervalo de confiança a 95% para a duração media do tempo de reparação e para a variância **[15]**

- b) Comente se é verdadeira ou falsa a afirmação: “Um intervalo de confiança a 95% para  $\theta$  contém, com probabilidade 0.95, o verdadeiro valor de  $\theta$ .” **[5]**

c) Com base na informação disponibilizada, estime a probabilidade de o tempo de reparação de uma máquina ser inferior a 100 minutos. **[10]**

d) Para um outro tipo de máquinas Y recolheu-se uma amostra aleatória de tempos de reparação (em minutos), com dados:  $n = 10$ ,  $\sum y_i = 986$ ,  $\sum y_i^2 = 99600$ . Determine com um grau de confiança 95% se o tempo medio de reparação é o mesmo para as duas máquinas. **[15]**

3. Um partido está a ponderar propôr uma candidata X para as eleições. Um grupo de membros não suporta esta candidata, pois considera que a proporção de votos a seu favor não seria superior a 25%. Para avaliar a suposição deste grupo, foi efetuado um inquérito a 1000 indivíduos, escolhidos ao acaso na população, tendo 270 deles manifestado a favor da candidata X.

a) Teste a suposição inicial do grupo do membros contra a candidata. Considere uma dimensão de 5% e de 10%. **[15]**

- b) Calcule o valor-p associado ao teste de hipóteses realizado em a), e interprete o seu significado estatístico e a informação adicional fornecida. **[5]**

4. Considera-se que o número de devoluções de artigos por dia em dada loja é uma variável aleatória com a seguinte distribuição de probabilidade:

$$f_X(X) = \begin{cases} 0.6, & X = 0 \\ 0.25, & X = 1 \\ 0.15, & X \geq 2 \end{cases}$$

Durante um período de 50 dias, houve 30 dias em que não se verificaram devoluções, houve 14 dias com uma devolução, e nos restantes dias houve duas ou mais devoluções.

Tendo em conta a amostra recolhida, analise a adequabilidade do modelo probabilístico considerado, realizando um teste de hipóteses ao nível de 5%. **[25]**

5. Considere o seguinte modelo linear

$$\text{salary}_i = \beta_0 + \beta_1 \text{years}_i + \beta_2 \text{gamesyr}_i + \beta_3 \text{bavg}_i + u_i$$

Observações:  $n = 353$  ( $i = 1, \dots, 353$ )

$y$ : salary	Salário do jogador de baseball
$x_1$ =years	Número de anos a jogar nas principais ligas
$x_2$ =gamesyr	Número de jogos por ano na liga
$x_3$ =bavg	Índice de performance no jogo

**MODELO:**  $\text{salary} \sim \text{years} + \text{gamesyr} + \text{bavg}$

Uma regressão OLS em R para este modelo, com todas as variáveis em escala linear, deu o resultado no Anexo I:

- a) Teste a significância estatística global da regressão e teste também a significância individual do parâmetro associado com a variável “years” para  $\alpha = 0.05$ . **[15]**
- b) Interprete as estimativas obtidas para os coeficientes associados com as variáveis “gamesyear” e “bavg”, indicando se estas variáveis são estatisticamente significativas. **[10]**
- c) Construa um intervalo de confiança a 90% para o coeficiente associado com a variável “bavg”. O que será possível concluir, a partir desse intervalo sobre a significância estatística do parâmetro? Justifique. **[15]**

- d) Qual o intervalo de previsão a 99% para o salário medio de um jogador com os seguintes parâmetros:

x1=5	Número de anos a jogar nas principais ligas
x2=110	Número de jogos por ano na liga
x3=215	Índice de performance no jogo

Use a regressão no Anexo II. Explique a que correspondem as variáveis explicativas desta regressão e construa o referido intervalo. **[15]**

Alterando a escala de y, para  $\log(\text{salário})$  obteve-se o resultado no Anexo III.

- e) Interprete o coeficiente associado a variável “years” desse modelo **[10]**.  
Explique os passos teóricos necessários para fazer uma previsão para y (*salário*) assumindo certos valores de (x1, x2, x3) a partir do modelo ajustado para  $\log(\text{salário})$ . **[10]**

**Anexo I**

lm(formula = salary ~ years + gamesyr + bavg)

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-1646063	389023	-4.231	2.97e-05 ***
years	64455	17771	3.627	0.000329 ***
gamesyr	19633	1974	9.946	< 2e-16 ***
bavg	3149	1566	2.011	0.045090 *

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1069000 on 349 degrees of freedom

Multiple R-squared: 0.4275,

Adjusted R-squared: 0.4226

F-statistic: 86.87 on 3 and 349 DF, p-value: &lt; 2.2e-16

**Anexo II**

lm(formula = salary ~ years1 + gamesyr1 + bavg1)

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1512888	111352	13.587	< 2e-16 ***
years1	64455	17771	3.627	0.000329 ***
gamesyr1	19633	1974	9.946	< 2e-16 ***
bavg1	3149	1566	2.011	0.045090 *

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1069000 on 349 degrees of freedom

Multiple R-squared: 0.4275,

Adjusted R-squared: 0.4226

F-statistic: 86.87 on 3 and 349 DF, p-value: &lt; 2.2e-16

**Anexo III**

lm(formula = Isalary ~ years + gamesyr + bavg)

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	10.910150	0.273602	39.876	< 2e-16 ***
years	0.070963	0.012498	5.678	2.87e-08 ***
gamesyr	0.019729	0.001388	14.210	< 2e-16 ***
bavg	0.001375	0.001101	1.248	0.213

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7521 on 349 degrees of freedom

Multiple R-squared: 0.5989,

Adjusted R-squared: 0.5954

F-statistic: 173.7 on 3 and 349 DF, p-value: &lt; 2.2e-16

