

**1º Semestre 2012/2013**
**2º Teste – Época de Recurso**

2013/02/01 – 15:00

**Justifique convenientemente todas as respostas!**

Duração: 90 min

**Grupo I**

10 valores

1. A duração, em dias, de uma componente electrónica é uma variável aleatória  $X$  com distribuição exponencial de parâmetro  $\lambda$ . Considere que  $(X_1, X_2, \dots, X_5)$  é uma amostra aleatória de tamanho 5 de  $X$ .

(a) Calcule a estimativa de máxima verosimilhança do parâmetro  $\lambda$  com base na seguinte concretização da amostra aleatória: 2.3, 2.7, 3.8, 4.3, 4.9. (3.0)

**Solução:**  $\hat{\lambda} = 0.278$ .

(b) Mostre que a variável aleatória  $Y = \min(X_1, X_2, \dots, X_5)$ , que representa o mínimo da amostra aleatória, tem distribuição exponencial de parâmetro  $5\lambda$ . (1.0)

(c) Tendo em conta o resultado da alínea anterior, investigue se  $5Y$  é ou não mais eficiente que a média amostral,  $\bar{X}$ , na estimação do parâmetro  $\mu = 1/\lambda$ . (2.0)

**Solução:**  $\bar{X}$  é mais eficiente que  $5Y$ .

2. Um investigador está interessado em comparar a altura de mulheres de dois países, A e B. Denotando por  $X_1$  e  $X_2$  as alturas, em centímetros, das mulheres nos países A e B, respectivamente, uma amostragem aleatória de 9 mulheres, de cada um dos países, produziu os seguintes valores:

País A	163	169	157	157	171	165	164	173	162	$\sum_{i=1}^9 x_{1i} = 1481$
País B	166	153	157	167	155	164	152	166	150	$\sum_{i=1}^9 x_{2i} = 1430$

Sabe-se que a variância das alturas das mulheres em cada um dos dois países é de  $40 \text{ cm}^2$ .

(a) Indicando hipóteses de trabalho que sejam necessárias, determine um intervalo de confiança a 99% para a diferença das alturas médias das mulheres nos países A e B. (3.0)

**Solução:** Admitindo  $X_i \sim \mathcal{N}(\mu, \sigma^2 = 40)$  e que ambas as amostras aleatórias são independentes,  $\text{IC}_{99\%}(\mu_1 - \mu_2) = [-2.013; 13.346]$ .

(b) Quantas mulheres devem ser inquiridas adicionalmente em cada país (supondo que os respectivos tamanhos amostrais se mantêm iguais) para estarmos confiante a 96% que a margem de erro na estimação da diferença das alturas médias referidas em (a) seja menor que 3 centímetros? (1.0)

**Solução:** 29.

**Grupo II**

10 valores

1. Escolhidos ao acaso 150 pagamentos registados numa loja, foram obtidos os seguintes dados: (3.5)

<i>Tipo de pagamento</i>	Numerário	Cheque	Cartão débito/crédito
<i>Número de pagamentos</i>	40	52	58

Teste a hipótese de as 3 modalidades de pagamento referidas serem equiprováveis. Calcule, justificando, o valor- $p$  do teste e decida com base no valor obtido, tendo em conta os níveis de significância usuais.

**Solução:** Valor- $p=0.186$ . Aos níveis de significância usuais (1%, 5%, 10%), não rejeito a hipótese de as 3 modalidades de pagamento serem equiprováveis.

2. Num certo dia de primavera, foram registados em 22 localidades de uma região de Inglaterra o número diário de horas de sol ( $x$ ) e a temperatura máxima diária ( $Y$ , em  $^{\circ}\text{C}$ ), estando os dados obtidos sumariados a seguir:

$$\sum_{i=1}^{22} x_i = 62.10 \quad \sum_{i=1}^{22} x_i^2 = 230.57 \quad \sum_{i=1}^{22} x_i y_i = 954.45 \quad \sum_{i=1}^{22} y_i = 331.00 \quad \sum_{i=1}^{22} y_i^2 = 4989.50$$

- (a) Ajuste um modelo de regressão linear simples de  $Y$  em função de  $x$ , indicando os pressupostos necessários para que esse modelo tenha validade estatística. (2.0)

**Solução:**  $\hat{Y} = 14.018 + 0.364x$ . Admite-se que os erros são variáveis aleatórias independentes, não correlacionadas entre si, com valor esperado nulo e variância constante.

- (b) Obtenha o coeficiente de determinação do modelo ajustado e comente o valor obtido. (1.0)

**Solução:**  $r^2 = 0.778$ . Cerca de 77.8% da variação total da temperatura máxima diária é explicada pelo número diário de horas de sol através do modelo de regressão considerado, valor suficientemente elevado para considerarmos que há um bom ajuste da recta estimada aos dados.

- (c) Teste, ao nível de significância de 1%, a significância do modelo de regressão linear. Comente o resultado do teste face ao valor obtido na alínea anterior. (3.5)

**Solução:** Ao nível de significância de 1%, rejeito a hipótese do coeficiente de regressão  $\beta_1$  ser nulo. Este resultado realça a importância do número diário de horas de sol na explicação da temperatura máxima diária, o que é coerente com a alínea anterior, onde se concluiu que há um bom ajuste da recta estimadas aos dados.