

Duração: **120** minutos

Exame de época normal

- As questões de escolha múltipla têm uma e apenas uma resposta correta. Caso assinale uma resposta errada ou mais de uma resposta, serão descontados 0.5 valores.
- As respostas às questões de escolha múltiplas têm de ser assinaladas na folha de instruções previamente distribuída.
- Para as restantes questões, terá de justificar convenientemente as suas respostas.
- Só pode sair da sala uma hora após o início do exame, devendo nesse caso entregar a sua prova ou desistir da mesma.

Pergunta 1

2 valores

Numa linha de produção admite-se que o número de defeitos em cada peça produzida se distribui de acordo com uma distribuição de Poisson com valor esperado igual a 0.1. Após a produção, todas as peças passam por uma linha de triagem onde todas as que não têm qualquer defeito, 80% das peças com um ou dois defeitos e 10% das peças com mais de dois defeitos são selecionadas para venda.

Calcule a probabilidade de uma peça não ter qualquer defeito, dado que foi selecionada para venda.

A: 0.8725 B: 0.9225 C: 0.7525 D: 0.9925

• **V.a. auxiliar, distribuição e f.p.**

X = número de defeitos numa peça produzida

$X \sim \text{Poisson}(\lambda)$, onde $\lambda : E(X) = 0.1 \Leftrightarrow \lambda = 0.1$.

$$P(X = x) \stackrel{\text{form.}}{=} e^{-0.1} \frac{0.1^x}{x!}, \quad x \in \mathbb{N}_0$$

• **Quadro de acontecimentos e probabilidades**

Acontecimento	Probabilidade
A = peça produzida sem defeitos	$P(A) = P(X = 0)$
B = peça produzida com um ou dois defeitos	$P(B) = P(X = 1) + P(X = 2)$
C = peça produzida com mais de dois defeitos	$P(C) = P(X > 2) = 1 - P(X \leq 2)$
V = peça selecionada para venda	
	$P(V A) = 1$
	$P(V B) = 0.8$
	$P(V C) = 0.1$

• **Prob. pedida**

Ao invocarmos o teorema de Bayes, temos

$$\begin{aligned}
 P(A | V) &= \frac{P(V | A) \times P(A)}{P(V)} \\
 &= \frac{P(V | A) \times P(A)}{P(V | A) \times P(A) + P(V | B) \times P(B) + P(V | C) \times P(C)} \\
 &= \frac{1 \times P(X = 0)}{1 \times P(X = 0) + 0.8 \times [P(X = 1) + P(X = 2)] + 0.1 \times [1 - P(X \leq 2)]} \\
 &= \frac{P(X = 0)}{1 \times P(X = 0) + 0.8 \times [P(X = 1) + P(X = 2)] + 0.1 \times [1 - P(X = 0) - P(X = 1) - P(X = 2)]}
 \end{aligned}$$

$$\begin{aligned}
 P(A|V) &= \frac{e^{-0.1}}{1 \times e^{-0.1} + 0.8 \times e^{-0.1} 0.1 + e^{-0.1} \frac{0.1^2}{2} + 0.1 \times \left(1 - e^{-0.1} - e^{-0.1} 0.1 - e^{-0.1} \frac{0.1^2}{2}\right)} \\
 &\approx \frac{0.904837}{0.980859} \\
 &\approx 0.9225.
 \end{aligned}$$

Pergunta 2

2 valores

Considere que o número de vendas semanais de equipamento de ultrassons profissional de uso terapêutico, pode ser considerado uma variável aleatória X com distribuição de Poisson com variância igual a 5 vendas semanais.

Determine $E(1.5X^2 - 0.3X + 2)$, o lucro esperado das vendas semanais.

- **V.a. de interesse**

X = número de vendas semanais de equipamento de ultrassons profissional de uso terapêutico

- **Distribuição de X**

$X \sim \text{Poisson}(\lambda)$, onde $\lambda : V(X) = E(X) = 5 \Leftrightarrow \lambda = 5$.

- **Valor esperado pedido**

$$\begin{aligned}
 E(1.5X^2 - 0.3X + 2) &= 1.5 \times E(X^2) - 0.3 \times E(X) + 2 \\
 &= 1.5 \times \{V(X) + [E(X)]^2\} - 0.3 \times E(X) + 2 \\
 &\stackrel{\text{form.}}{=} 1.5 \times (5 + 5^2) - 0.3 \times 5 + 2 \\
 &= 1.5 \times 30 - 0.3 \times 5 + 2 \\
 &= 45.5.
 \end{aligned}$$

Pergunta 3

2 valores

Considere que a velocidade do vento numa certa região (em dezenas de km/h) é uma variável aleatória X com função de densidade de probabilidade $f_X(x) = (x/9) \times e^{-x^2/18}$, para $x \geq 0$.

No planeamento de uma estrutura é importante levar em conta a possibilidade de ocorrerem condições meteorológicas extremas. Nesse sentido, calcule o valor do quantil de probabilidade 0.999 de X em km/h.

- **V.a. de interesse e f.d.p.**

X = velocidade do vento numa certa região (em dezenas de km/h)

$$f_X(x) \stackrel{\text{form.}}{=} \begin{cases} \frac{x}{9} \times e^{-\frac{x^2}{18}}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

- **Quantil pedido**

Pretende-se determinar

$$\begin{aligned}
 q \times 10 : P(X \leq q) = 0.999 &\Leftrightarrow \int_{-\infty}^q f_X(x) dx = 0.999 \Leftrightarrow \int_0^q \frac{x}{9} e^{-\frac{x^2}{18}} dx = 0.999 \Leftrightarrow -e^{-\frac{x^2}{18}} \Big|_0^q = 0.999 \\
 1 - e^{-\frac{q^2}{18}} = 0.999 &\Leftrightarrow q = \sqrt{-18 \ln 0.001} \approx 11.15.
 \end{aligned}$$

Assim sendo, $q \times 10 = 111.5 \text{ km/h}$.

Uma loja de comércio eletrônico vende um dado equipamento em dois estados: novo ou reconcondicionado. Para uma qualquer venda desse equipamento considere as variáveis aleatórias X , que indica o estado do produto (0: reconcondicionado e 1: novo), e Y que representa o número de trocas no primeiro mês após a venda. Considere que a função de probabilidade conjunta de (X, Y) está descrita abreviadamente na tabela à direita.

X	Y		
	0	1	2
0	0.1	0.1	0.2
1	0.4	0.2	0.0

Calcule o valor do coeficiente de correlação entre as variáveis X e Y .

A: -0.5750 B: 0.575 C: 0.4250 D: -0.1050

• **Par aleatório**

(X, Y)

X = estado do produto (0: reconcondicionado e 1: novo)

Y = número de trocas no primeiro mês após a venda

• **Correlação pedida**

Uma vez que $corr(X, Y) = \frac{cov(X, Y)}{\sqrt{V(X) \times V(Y)}}$, são necessários alguns cálculos auxiliares que envolverão a f.p. conjunta de (X, Y) e as f.p. marginais de X e Y dadas por $P(X = x) = \sum_y P(X = x, Y = y)$ e $P(Y = y) = \sum_x P(X = x, Y = y)$.

X	Y			$P(X = x)$
	0	1	2	
0	0.1	0.1	0.2	0.4
1	0.4	0.2	0.0	0.6
$P(Y = y)$	0.5	0.3	0.2	1

• **Valores esperados e variâncias de X e de Y**

$$E(X) = \sum_{x=0}^1 x \times P(X = x) = 1 \times 0.6 = 0.6$$

$$E(X^2) = \sum_{x=0}^1 x^2 P(X = x) = 1^2 \times 0.6 = 0.6$$

$$V(X) = E(X^2) - E^2(X) = 0.6 - 0.6^2 = 0.24$$

$$E(Y) = \sum_{y=0}^2 y \times P(Y = y) = 1 \times 0.3 + 2 \times 0.2 = 0.7$$

$$E(Y^2) = \sum_{y=0}^2 y^2 P(Y = y) = 1^2 \times 0.6 + 2^2 \times 0.1 = 1.1$$

$$V(Y) = E(Y^2) - E^2(Y) = 1.1 - 0.7^2 = 0.61$$

• **Valor esperado de XY e covariância entre X e Y**

$$E(XY) = \sum_{x=0}^1 \sum_{y=0}^2 xy \times P(X = x, Y = y) = 1 \times 1 \times 0.2 = 0.2$$

$$cov(X, Y) = E(XY) - E(X) \times E(Y) = 0.2 - 0.6 \times 0.7 = -0.22$$

• **Valor esperado e variância pedidos (cont.)**

$$\begin{aligned} corr(X, Y) &= \frac{cov(X, Y)}{\sqrt{V(X) \times V(Y)}} \\ &= \frac{-0.22}{\sqrt{0.24 \times 0.61}} \\ &\approx -0.5750. \end{aligned}$$

Numa experiência laboratorial admite-se que cada partícula emitida é detetada por um sensor com uma probabilidade de 0.95. A experiência é realizada em três etapas independentes em que são emitidas 3, 7 e 10 partículas em cada uma dessas etapas. Seja Y a variável aleatória que representa o número de partículas detetadas no conjunto das três etapas da experiência.

Obtenha a probabilidade de Y não exceder 18 sabendo que foi detetada pelo menos uma partícula no conjunto das três etapas da experiência.

- **V.a. e distribuições**

X_i = número de partículas detetadas pelo sensor na etapa i , $i = 1, 2, 3$

$X_i \stackrel{indep.}{\sim} \text{binomial}(n_i, p)$, $n_1 = 3$, $n_2 = 7$, $n_3 = 10$, $p = 0.95$

- **V.a. de interesse**

$Y = \sum_{i=1}^3 X_i$ = número de partículas detetadas no conjunto das três etapas da experiência

- **Distribuição exacta de Y**

$Y \sim \text{binomial}(n = \sum_{i=1}^3 n_i = 20, p = 0.95)$, por admitirmos independência entre as três etapas da experiência.

- **F.p. de Y**

$P(Y = y) = \binom{20}{y} 0.95^y (1 - 0.95)^{20-y}$, $y = 0, 1, \dots, 20$

- **Prob. pedida**

$$P(Y \leq 18 | Y > 0) = \frac{P(Y \leq 18, Y > 0)}{P(Y > 0)} = \frac{P(0 < Y \leq 18)}{1 - P(Y \leq 0)} = \frac{P(Y \leq 18) - P(Y \leq 0)}{1 - P(Y \leq 0)} = \frac{F_Y(18) - F_Y(0)}{1 - F_Y(0)}$$

$\stackrel{calc.}{=} 0.2642.$

[Alternativamente, pode recorrer-se aos resultados $Z = n - Y \sim \text{binomial}(n = 20, 1 - p = 0.05)$ e $F_Y(y) = P(Y \leq y) = P(Z = n - Y \geq n - y) = 1 - P(Z \leq n - y - 1) = 1 - F_Z(n - y - 1)$ e concluir que

$$P(Y \leq 18 | Y > 0) = \frac{F_Y(18) - F_Y(0)}{1 - F_Y(0)} = \frac{[1 - F_Z(20 - 18 - 1)] - [1 - F_Z(20 - 0 - 1)]}{1 - [1 - F_Z(20 - 0 - 1)]} = \frac{F_Z(19) - F_Z(1)}{F_Z(19)}$$

$$= 1 - \frac{F_Z(1)}{F_Z(19)} \stackrel{tabelas}{\approx} 1 - \frac{0.7358}{1.0000} = 0.2642.]$$

Admita que a proporção de combustível no depósito de automóveis acidentados é uma variável aleatória X com função de densidade de probabilidade $f_X(x) = 2\theta x(1 - x^2)^{\theta-1}$, para $0 < x < 1$ e $\theta \in \mathbb{R}^+$. Uma amostra referente a 100 acidentes, $(x_1, \dots, x_{100})$, conduziu ao valor $\sum_{i=1}^{100} \ln(1 - x_i^2) \approx -45.87$.

Com base nos dados observados, calcule a estimativa de máxima verosimilhança da mediana de X , $F_X^{-1}(0.5) = \sqrt{1 - 0.5^{1/\theta}}$. **A:** 0.8828 **B:** 0.6860 **C:** 0.5219 **D:** 0.8530

- **V.a. de interesse; f.d.p.**

X = a proporção de combustível no depósito de automóveis acidentados

$$f_X(x) = \begin{cases} 2\theta x(1 - x^2)^{\theta-1}, & 0 < x < 1 \\ 0, & \text{outros valores de } x \end{cases}$$

- **Parâmetro desconhecido e espaço paramétrico**

θ , $\Theta = \mathbb{R}^+$

- **Amostra**

$(x_1, \dots, x_n)$, amostra de dimensão n proveniente de X .

- **Dedução da estimativa de MV de θ**

Passo 1 — Função de verosimilhança

$$\begin{aligned} L(\theta | \underline{x}) &= f_{\underline{X}}(\underline{x}) \stackrel{X_i \text{ indep}}{=} \prod_{i=1}^n f_{X_i}(x_i) \stackrel{X_i \sim X}{=} \prod_{i=1}^n f_X(x_i) \\ &= \prod_{i=1}^n \left[2\theta x_i (1 - x_i^2)^{\theta-1} \right] \\ &= 2^n \theta^n \prod_{i=1}^n \left[x_i (1 - x_i^2)^{\theta-1} \right], \quad \theta > 0 \end{aligned}$$

Passo 2 — Função de log-verosimilhança

$$\ln L(\theta | \underline{x}) = n \ln(2) + n \ln(\theta) + \sum_{i=1}^n \ln(x_i) + (\theta - 1) \sum_{i=1}^n \ln(1 - x_i^2), \quad \theta > 0$$

Passo 3 — Maximização

A estimativa de MV de θ será representada por $\hat{\theta}$ e

$$\hat{\theta} : \begin{cases} \left. \frac{d \ln L(\theta | \underline{x})}{d\theta} \right|_{\theta=\hat{\theta}} = 0 & \text{(ponto de estacionaridade)} \\ \left. \frac{d^2 \ln L(\theta | \underline{x})}{d\theta^2} \right|_{\theta=\hat{\theta}} < 0 & \text{(ponto de máximo)} \\ \left\{ \begin{array}{l} \frac{n}{\hat{\theta}} + \sum_{i=1}^n \ln(1 - x_i^2) = 0 \\ -\frac{n}{\hat{\theta}^2} < 0 \quad \text{(prop. verdadeira)} \end{array} \right. \end{cases}$$

Passo 4 — Estimativa de MV de θ

$$\hat{\theta} = -\frac{n}{\sum_{i=1}^n \ln(1 - x_i^2)} \simeq -\frac{100}{-45.87} = \frac{1}{0.4587}$$

- **Outro parâmetro desconhecido**

$$h(\theta) = me(X) = \sqrt{1 - 0.5^{1/\theta}}$$

$$[\text{Obs.: } q = me(X) : P(X \leq q) = 0.5 \Leftrightarrow \int_{-\infty}^q f_X(x) dx = 0.5 \Leftrightarrow \int_0^q 2\theta x(1 - x^2)^{\theta-1} dx = 0.5 \Leftrightarrow 1 - (1 - q^2)^\theta = 0.5 \Leftrightarrow q = \sqrt{1 - 0.5^{1/\theta}}.]$$

- **Estimativa de MV de $h(\theta)$**

Ao invocar a propriedade de invariância dos EMV, a estimativa de MV de $h(\theta)$ é dada por

$$\widehat{h(\theta)} = h(\hat{\theta}) = \sqrt{1 - 0.5^{1/\hat{\theta}}} = \sqrt{1 - 0.5^{1/(1/0.4587)}} = \sqrt{1 - 0.5^{0.4587}} \simeq 0.5219.$$

Pergunta 7

2 valores

Uma engenheira civil estuda a resistência à compressão de um material cerâmico e tem dúvidas acerca do valor esperado da resistência anunciada pelo fabricante do material. A soma de 40 valores da resistência obtidos em ensaios laboratoriais foi igual a 1365 MPa (megapascals).

Admitindo que o desvio padrão da resistência à compressão do material é igual a 4 MPa, calcule um intervalo de confiança para o respetivo valor esperado a um nível de confiança aproximado de 0.96.

- **V.a. de interesse**

X = resistência à compressão de um material cerâmico

- **Situação**

X com distribuição arbitrária

$\mu = E(X)$ DESCONHECIDO

$\sigma^2 = V(X) = 4^2$ conhecida

$n = 40 \geq 30$

- **Obtenção de IC aproximado para μ**

Passo 1 - Seleção da v.a. fulcral para μ

Pelo TLC temos

$$Z = \frac{\bar{X} - E(X)}{\sqrt{\frac{V(X)}{n}}} = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \stackrel{a}{\sim} \text{normal}(0, 1).$$

Passo 2 — Obtenção dos quantis de probabilidade

Como $(1 - \alpha) \times 100\% = 96\% \Leftrightarrow \alpha = 0.04$, lidaremos com os quantis seguintes:

$$\begin{cases} a_\alpha = \Phi^{-1}(\alpha/2) = -\Phi^{-1}(1 - \alpha/2) = -\Phi^{-1}(0.98) \stackrel{\text{tabela/calc.}}{=} -2.0537 \\ b_\alpha = \Phi^{-1}(1 - \alpha/2) = \Phi^{-1}(0.98) = 2.0537. \end{cases}$$

Passo 3 — Inversão da desigualdade $a_\alpha \leq Z \leq b_\alpha$

$$P(a_\alpha \leq Z \leq b_\alpha) \simeq 1 - \alpha$$

$$P\left[a_\alpha \leq \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \leq b_\alpha\right] \simeq 1 - \alpha$$

$$P\left[\bar{X} - b_\alpha \times \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X} - a_\alpha \times \frac{\sigma}{\sqrt{n}}\right] \simeq 1 - \alpha$$

Passo 4 — Concretização

A expressão geral do intervalo aproximado de confiança a $(1 - \alpha) \times 100\%$ para μ é

$$IC_{(1-\alpha) \times 100\%}(\mu) \simeq \left[\bar{x} \pm \Phi^{-1}(1 - \alpha/2) \times \frac{\sigma}{\sqrt{n}} \right].$$

Dado que $n = 40$, $\alpha = 0.04$, $\Phi^{-1}(1 - \alpha/2) = 2.0537$ e $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1365}{40} = 34.125$,

$$\begin{aligned} IC_{96\%}(\mu) &\simeq \left[34.125 - 2.0537 \times \frac{4}{\sqrt{40}}, 34.125 + 2.0537 \times \frac{4}{\sqrt{40}} \right] \\ &\simeq [32.8261, 35.4239]. \end{aligned}$$

Pergunta 8

2 valores

Julga-se que a prevalência da deuteranopia (uma forma de daltonismo) numa certa população masculina é de 3%. Um investigador deseja testar esta hipótese e selecionou ao acaso 1200 homens do registo dos pacientes de um grande hospital tendo constatado que 22 deles sofrem de deuteranopia.

Teste a hipótese $H_0 : p = 0.03$ contra $H_1 : p \neq 0.03$ a um nível de significância de 5%.

- **V.a. de interesse**

$$X = \begin{cases} 1, & \text{se o paciente sofre de deuteranopia} \\ 0, & \text{c.c.} \end{cases}$$

- **Situação**

$X \sim \text{Bernoulli}(p)$

$p = P(\text{paciente sofre de deuteranopia})$ DESCONHECIDA

$n = 1200 \geq 30$

- **Hipóteses**

$H_0 : p = p_0 = 0.03$

$H_1 : p \neq p_0$

- **N.s.**

$\alpha_0 = 5\%$

- **Estatística de teste**

$$T = \frac{\bar{X} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \stackrel{a}{\sim}_{H_0} \text{normal}(0, 1)$$

- **Região de rejeição de H_0** (para valores de T)

Lidamos com um teste bilateral ($H_1 : p \neq p_0$), logo a região de rejeição de H_0 é do tipo $W = (-\infty, -c) \cup (c, +\infty)$, onde

$$c = \phi^{-1}(1 - \alpha_0/2) = \phi^{-1}(1 - 0.05/2) = \phi^{-1}(0.975) \stackrel{\text{tabela, calc.}}{=} 1.9600.$$

Logo, $P(\text{Rejeitar } H_0 \mid H_0) \simeq \alpha_0$.

- **Decisão**

O valor observado da estatística de teste é

$$t = \frac{\bar{x} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\frac{22}{1200} - 0.03}{\sqrt{\frac{0.03 \times (1-0.03)}{1200}}} \simeq -2.37.$$

Como $t \simeq -2.37 \in W = (-\infty, -1.9600) \cup (1.9600, +\infty)$, devemos rejeitar H_0 ao n.s. $\alpha_0 = 5\%$ [e a qualquer outro n.s. superior a $\alpha_0 = 5\%$.]

Pergunta 9

2 valores

O processo de enchimento de latas de silicone líquido está automatizado em determinada unidade fabril. A gestora desta unidade conjectura que a massa (em kg) de silicone em cada lata (X) possui distribuição normal com valor esperado 25 e desvio padrão unitário (hipótese H_0). Para avaliar esta conjectura, recolheu-se uma amostra casual de X , tendo-se agrupado as 100 observações e obtido a tabela seguinte com as frequências observadas a que juntou as frequências esperadas sob H_0 :

X	≤ 23.5	$]23.5, 24.5]$	$]24.5, 25.5]$	$]25.5, 26.5]$	> 26.5
Frequência absoluta observada	10	28	42	19	1
Frequência absoluta esperada sob H_0	6.681	24.173	38.292	24.173	6.681

Com base no teste de ajustamento do qui-quadrado, a gestora deve

A: Rejeitar H_0 a 1%, 5% e 10%.

B: Rejeitar H_0 a 5% e 10% e não rejeitar H_0 a 1%.

C: Rejeitar H_0 a 10% e não rejeitar H_0 a 1% e 5%.

D: Não rejeitar H_0 a 1%, 5% e 10%.

- **V.a. de interesse**

X = massa (em kg) de silicone em cada lata

- **Hipóteses**

$H_0 : X \sim \text{normal}(\mu = 25, \sigma^2 = 1)$ vs. $H_1 : X \not\sim \text{normal}(\mu = 25, \sigma^2 = 1)$

- **Estatística de teste**

$$T = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \underset{H_0}{\sim} \chi^2_{(k-1)}$$

onde: k = número de classes = 5; O_i = frequência absoluta observável da classe i ; E_i = frequência absoluta esperada, sob H_0 , da classe i .

- **Região de rejeição de H_0** (para valores de T)

A região de rejeição de H_0 dum teste de ajustamento do qui-quadrado é um intervalo à direita $W = (c, +\infty)$.

- **Decisão (com base no valor-p aproximado)**

i	Classe	Freq. abs. obs. o_i	Freq. abs. esper. sob H_0 $E_i = n \times p_i^0$	Parcelas valor obs. estat. teste $\frac{(o_i - E_i)^2}{E_i}$
1	(0, 23.5]	10	6.681	$\frac{(10 - 6.681)^2}{6.681} = 1.649$
2	(23.5, 24.5]	28	24.173	0.606
3	(24.5, 25.5]	42	38.292	0.359
4	(25.5, 26.5]	19	24.173	1.107
5	(26.5, $+\infty$)	1	6.681	4.831
		$\sum_{i=1}^k o_i = n = 100$	$\sum_{i=1}^k E_i = n = 100$	$t = \sum_{i=1}^k \frac{(o_i - E_i)^2}{E_i} = 8.552$

Atendendo a que a região de rejeição de H_0 é um intervalo à direita num teste de ajustamento do qui-quadrado, temos

$$\text{valor} - p = P(T > t | H_0) \simeq 1 - F_{\chi^2_{(5-1)}}(8.552) \stackrel{\text{calc.}}{\simeq} 0.0733.$$

Assim sendo, devemos decidir pela:

- rejeição de H_0 a qualquer n.s. $\alpha_0 > 7.33\%$, designadamente ao n.u.s. de 10%;
- não rejeição de H_0 a qualquer n.s. $\alpha_0 \leq 7.33\%$, nomeadamente ao n.u.s. de 1% e 5%.

Por consequência, a resposta, das quatro apresentadas, que se coaduna o referido acima é:

- **C:** Rejeitar H_0 a 10% e não rejeitar H_0 a 1% e 5%.

[Alternativamente, a consulta das tabelas dos quantis de probabilidade da distribuição do qui-quadrado com $k - 1 = 4$ graus de liberdade permitem concluir que

$$\begin{aligned} F_{\chi^2_{(5-1)}}^{-1}(0.9) = 7.779 &< 8.552 < 8.496 = F_{\chi^2_{(5-1)}}^{-1}(0.925) \\ 0.075 = 1 - 0.925 &< 1 - F_{\chi^2_{(5-1)}}(8.552) < 1 - 0.90 = 0.10. \end{aligned}$$

Logo, devemos: rejeitar H_0 a qualquer n.s. $\alpha_0 \geq 10\%$, designadamente ao n.u.s. de 10%; não rejeitar de H_0 a qualquer n.s. $\alpha_0 \leq 7.5\%$, nomeadamente aos n.u.s. de 1% e 5%. Deste modo, a resposta correcta é **C:** Rejeitar H_0 a 10% e não rejeitar H_0 a 1% e 5%.]

Pergunta 10

2 valores

Pretende-se investigar a relação entre a massa de cada paciente (x , em quilos) e a distância percorrida no teste dos 6 minutos de marcha (Y , em metros). Numa amostra de 5 pacientes de um hospital, obtiveram-se os seguintes resultados:

$$\sum_{i=1}^5 x_i = 466, \quad \sum_{i=1}^5 x_i^2 = 43482, \quad \sum_{i=1}^5 y_i = 1035, \quad \sum_{i=1}^5 y_i^2 = 215425, \quad \sum_{i=1}^5 x_i y_i = 96320,$$

onde $[\min_{i=1, \dots, 5} x_i, \max_{i=1, \dots, 5} x_i] = [87, 96]$. Admita que x e Y estão relacionadas de acordo com o modelo de regressão linear simples, $Y = \beta_0 + \beta_1 x + \epsilon$.

Estime $E(Y | x_0 = 89)$; calcule o coeficiente de determinação e interprete o seu valor.

- **[Modelo de RLS e hipóteses de trabalho]**

Y = distância percorrida no teste dos 6 minutos de marcha em metros (v.a. resposta)

x = massa do paciente em kilos (variável explicativa)

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, \dots, n$$

$$E(\varepsilon_i) = 0, \quad V(\varepsilon_i) = \sigma^2, \quad i = 1, \dots, n$$

$$\text{cov}(\varepsilon_i, \varepsilon_j) = 0, \quad i \neq j, \quad i, j = 1, \dots, n$$

- **Estimativas de MQ dos parâmetros desconhecidos β_0 e β_1**

Temos

- $n = 5$

- $\sum_{i=1}^n x_i = 466$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{466}{5} = \frac{466}{5} = 93.2$$

$$\sum_{i=1}^n x_i^2 = 43482$$

$$\sum_{i=1}^n x_i^2 - n \bar{x}^2 = 43482 - 5 \times 93.2^2 = 50.8$$

- $\sum_{i=1}^n y_i = 1035$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{1035}{5} = 207$$

$$\sum_{i=1}^n y_i^2 = 215425$$

$$\sum_{i=1}^n y_i^2 - n \bar{y}^2 = 215425 - 5 \times 207^2 = 1180$$

- $\sum_{i=1}^n x_i y_i = 96320$

$$\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y} = 96320 - 5 \times 93.2 \times 207 = -142.$$

Logo, as estimativas de MQ dos parâmetros desconhecidos β_1 e β_0 são dadas por

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sum_{i=1}^n x_i^2 - n \bar{x}^2} = -\frac{142}{50.8} \simeq -2.795276$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \simeq 207 - (-2.795276) \times 93.2 \simeq 467.519685.$$

- **Estimativa pedida**

$$\begin{aligned} \hat{E}(Y | x_0 = 89) &= \hat{\beta}_0 + \hat{\beta}_1 \times x_0 \\ &\simeq 467.519685 - 2.795276 \times 89 \\ &\simeq 218.740121. \end{aligned}$$

[Uma vez que $x_0 = 89 \in [\min_{i=1, \dots, 5} x_i, \max_{i=1, \dots, 5} x_i] = [87, 96]$ não cometemos qualquer erro de extrapolação.]

- **Coeficiente de determinação**

$$\begin{aligned} r^2 &= \frac{(\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y})^2}{(\sum_{i=1}^n x_i^2 - n \bar{x}^2) \times (\sum_{i=1}^n y_i^2 - n \bar{y}^2)} \\ &= \frac{(-142)^2}{50.8 \times 1180} \\ &\simeq 0.336381. \end{aligned}$$

- **Interpretação do coeficiente de determinação**

Cerca de 34% da variação total da variável resposta Y é explicada pela variável x , através do modelo de regressão linear simples ajustado. Podemos concluir que a recta estimada não se ajusta bem ao conjunto de dados.