

Duração: **120** minutos

Exame de Época Normal — C

- 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

Um teste de deteção de um parasita numa dada população dá resultado positivo (i.e., indica que o indivíduo é hospedeiro do parasita) com probabilidade: 0.85, se o indivíduo testado for hospedeiro do parasita; e 0.05, se o indivíduo testado não for hospedeiro do parasita. Sabe-se ainda que o teste dá resultado positivo em 25% dos indivíduos da população.

Selecionado ao acaso um indivíduo da população, o valor mais aproximado da probabilidade do indivíduo ser hospedeiro do parasita é

A: 0.24 **B:** 0.25 **C:** 0.27 **D:** 0.28

- **Quadro de acontecimentos e probabilidades para um indivíduo escolhido ao acaso da população**

Acontecimento	Probabilidade
H = indivíduo é hospedeiro do parasita	$P(H) = x$
T = teste aplicado ao indivíduo dá resultado positivo	$P(T) = 0.25$
	$P(T H) = 0.85$
	$P(T \bar{H}) = 0.05$

- **Prob. pedida**

Ao invocarmos o teorema da probabilidade total, temos

$$P(T) = P(T | H) \times P(H) + P(T | \bar{H}) \times P(\bar{H})$$

$$\Leftrightarrow 0.25 = 0.85 \times x + 0.05 \times (1 - x)$$

$$\Leftrightarrow 0.20 = 0.80x$$

$$\Leftrightarrow x = \frac{0.20}{0.80} = 0.25.$$

Pergunta 2

2 valores

Sabe-se que em certa região, a proporção de habitantes com o tipo sanguíneo B é igual a $\frac{1}{8}$. Sejam: X a variável aleatória que representa o número de dadores voluntários que se dirigem a um serviço de recolha de sangue até surgir o primeiro dador com o tipo sanguíneo B ; e ξ o primeiro quartil de X .

Calcule a probabilidade de X não exceder o dobro de ξ dado que X já é superior a ξ .

- **V.a.**

X = no. de dadores voluntários que se dirigem a [...] até surgir o 1o. dador com o tipo sanguíneo B ...

$X \sim \text{geométrica}(p = \frac{1}{8})$

- **F.p. e f.d. de X**

$$P(X = x) = (1 - p)^{x-1} p = \left(1 - \frac{1}{8}\right)^{x-1} \times \frac{1}{8}, \quad x \in \mathbb{N}$$

$$F_X(x) = P(X \leq x) = \sum_{n=1}^x P(X = n) = 1 - (1 - p)^x = 1 - \left(1 - \frac{1}{8}\right)^x = 1 - 0.875^x, \quad x \in \mathbb{N}$$

- **Quantil pedido**

O quantil de probabilidade 0.25 de X , ξ , define-se do modo seguinte:

$$\xi \in \mathbb{N} : P(X \leq \xi) \geq 0.25 \quad \text{e} \quad P(X \geq \xi) \geq 1 - 0.25 \quad \Leftrightarrow \quad 0.25 \leq F_X(\xi) \leq 0.25 + P(X = \xi)$$

$$F_X[(\xi)^-] \leq 0.25 \leq F_X(\xi).$$

Ao recorrermos à f.d. acima e a uma calculadora, temos $F_X(1) = 1 - 0.875 = 0.125$, $F_X(2) = F_X(3^-) = 1 - 0.875^2 \approx 0.234$ e $F_X(3) = 1 - 0.875^3 \approx 0.330$. Logo,

$$0.234 \approx F_X(3^-) \leq 0.25 \leq F_X(3) \approx 0.330,$$

pelo que o primeiro quantil de X é igual a $\xi = 3$.

- **Prob. pedida**

Ao invocarmos a propriedade de falta de memória da distribuição geométrica, temos

$$\begin{aligned} P(X \leq 2\xi \mid X > \xi) &= [1 - P(X > 2\xi \mid X > \xi)] = 1 - P(X > 2\xi - \xi) \\ &= P(X \leq 2\xi - \xi) = F_X(\xi) \stackrel{\xi=3}{\approx} 0.330. \end{aligned}$$

[Alternativamente,

$$\begin{aligned} P(X \leq 2\xi \mid X > \xi) &= \frac{P(X \leq 2\xi, X > \xi)}{P(X > \xi)} = \frac{P(\xi < X \leq 2\xi)}{1 - P(X \leq \xi)} = \frac{F_X(2\xi) - F_X(\xi)}{1 - F_X(\xi)} \\ &= \frac{[1 - (1 - p)^{2\xi}] - [1 - (1 - p)^\xi]}{(1 - p)^\xi} = \frac{(1 - p)^\xi [1 - (1 - p)^\xi]}{(1 - p)^\xi} = 1 - (1 - p)^\xi \\ &= F_X(\xi) = 1 - 0.875^3 \approx 0.330.] \end{aligned}$$

Pergunta 3

2 valores

Uma fábrica produz peças industriais e a variável aleatória X (em horas), que indica a diferença entre o tempo de produção efetivo de uma peça e o tempo ideal, possui função de densidade de probabilidade simétrica dada por

$$f_X(x) = \begin{cases} \frac{4 - |x|}{16}, & -4 < x < 4, \\ 0, & \text{caso contrário.} \end{cases}$$

O custo adicional associado a atrasos ou adiantamentos no tempo de produção de uma peça é $Y = 3X + 5X^2$.

Obtenha o valor esperado de Y .

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

X = tempo de produção de uma peça

$$f_X(x) = \frac{4 - |x|}{16}, \quad -4 < x < 4$$

- **V.a. de interesse**

$Y = 3X + 5X^2$ = custo adicional, assoc. a atrasos/adiantamentos no tempo de produção de uma peça

- **Valor esperado pedido**

$E(Y) = E(3X + 5X^2) = 3E(X) + 5E(X^2)$, onde:

$$E(X) = \int_{-\infty}^{+\infty} x f_X(x) dx = \frac{1}{16} \int_{-4}^4 x(4 - |x|) dx = \frac{1}{16} \left[\int_{-4}^0 x(4 + x) dx + \int_0^4 x(4 - x) dx \right]$$

$$E(X) = \frac{1}{16} \left(2x^2 + \frac{x^3}{3} \right) \Big|_{-4}^0 + \frac{1}{16} \left(2x^2 - \frac{x^3}{3} \right) \Big|_0^4 = 0$$

[alternativamente, poderíamos ter invocado o facto de a f.d.p. ser simétrica em torno da origem e concluir que $E(X) = 0$];

$$\begin{aligned} E(X^2) &= \int_{-\infty}^{+\infty} x^2 f_X(x) dx = \frac{1}{16} \int_{-4}^4 x^2 (4 - |x|) dx = \frac{1}{16} \left[\int_{-4}^0 x^2 (4 + x) dx + \int_0^4 x^2 (4 - x) dx \right] \\ &= \frac{1}{16} \left(\frac{4x^3}{3} + \frac{x^4}{4} \right) \Big|_{-4}^0 + \frac{1}{16} \left(\frac{4x^3}{3} - \frac{x^4}{4} \right) \Big|_0^4 = \frac{1}{16} \times 2 \times \left(\frac{4x^4}{3} - \frac{x^4}{4} \right) \Big|_0^4 = \frac{1}{8} \times \frac{64}{3} = \frac{8}{3}. \end{aligned}$$

Consequentemente,

$$E(Y) = 3E(X) + 5E(X^2) = 5E(X^2) = \frac{40}{3} = 13.(3).$$

Pergunta 4

2 valores

Seja X a variável aleatória que representa a diferença (em dezenas de Euros) entre o montante mensal efectivamente despendido numa oficina metalo-mecânica e o montante orçamentado. Defina-se agora a variável aleatória $Y = -X$, que facilita a visualização da margem de segurança ou do excesso de despesa.

Deduzo o valor da correlação entre X e Y e comente o resultado.

- **V.a. e f.p.**

X = diferença entre o montante despendido mensalmente e o montante orçamentado

$Y = -X$ = diferença entre o montante orçamentado e o montante despendido mensalmente

- **Correlação pedida**

$$\begin{aligned} \text{corr}(X, Y) &= \frac{\text{cov}(X, Y)}{\sqrt{V(X)} \sqrt{V(Y)}} = \frac{E(XY) - E(X) \times E(Y)}{\sqrt{V(X)} \sqrt{V(Y)}} = \frac{E[X(-X)] - E(X) \times E(-X)}{\sqrt{V(X)} \sqrt{V(-X)}} \\ &= \frac{-E(X^2) + E^2(X)}{\sqrt{V^2(X)}} = \frac{-V(X)}{V(X)} = -1. \end{aligned}$$

- **Comentário pedido**

O resultado $\text{corr}(X, Y) = -1$ indica que X e Y são linear/perfeita e negativamente correlacionadas. Em particular, sempre que X assume um determinado valor, Y assume exactamente o valor simétrico.

Pergunta 5

2 valores

Num jogo eletrónico competem equipas com um qualquer número de jogadores. Uma das tarefas de cada jogador é a recolha de artefatos que se encontram, em grande número, escondidos no cenário do jogo. Admita que o número de artefatos que são recolhidos por um jogador em cada hora de jogo é uma variável aleatória com distribuição de Poisson de valor esperado 0.5 e independente do número de artefatos recolhidos por qualquer outro jogador.

A dimensão mínima que uma equipa deve ter para se garantir que a probabilidade da equipa não recolher qualquer artefato durante uma hora não excede 0.05 é

A: 6 B: 2 C: 5 D: 7

- **V.a.**

A_i = número de artefatos recolhidos pelo jogador i , $i = 1, \dots, n$

$A_i \stackrel{\text{indep.}}{\sim} \text{Poisson}(0.5)$

- **V.a. de interesse**

$A = \sum_{i=1}^n A_i$ = número total de artefactos recolhidos por uma equipa com n jogadores

$A \sim \text{Poisson}(0.5n)$

$$P(A = x) = e^{-0.5n} \frac{(0.5n)^x}{x!}, \quad x \in \mathbb{N}_0$$

- **Dimensão mínima pedida**

$$n : P(A = 0) \leq 0.05 \Leftrightarrow e^{-0.5n} \leq 0.05 \Leftrightarrow -0.5n \leq \ln(0.05) \Leftrightarrow n \geq -\frac{1}{0.5} \ln(0.05) \approx 5.991465.$$

Logo, o tamanho mínimo da equipa deverá ser **A: 6**.

Pergunta 6

2 valores

Admita que a taxa de transferência de dados num nó de uma rede de computadores (em GB/s) é uma variável aleatória X com função de densidade de probabilidade dada por

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{[\ln(x)-2]^2}{2\sigma^2}\right\}, \quad x > 0,$$

onde σ denota um parâmetro desconhecido positivo. Uma amostra referente a $n = 100$ medições dessa taxa, $(x_1, \dots, x_n)$, conduziu ao valor $\sum_{i=1}^n [\ln(x_i) - 2]^2 \approx 1.0233$.

Com base nestes dados, a estimativa de máxima verosimilhança de $E(X) = \exp(2 + \sigma^2/2)$ é aproximadamente igual a **A: 7.43 B: 7.77 C: 12.33 D: 6.66**

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

X = taxa de transferência de dados num nó de uma rede de computadores

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left\{-\frac{[\ln(x)-2]^2}{2\sigma^2}\right\}, \quad x > 0$$

- **Parâmetro desconhecido**

σ ($\sigma > 0$)

- **Amostra**

$(x_1, \dots, x_n) : n = 100, \quad \sum_{i=1}^n [\ln(x_i) - 2]^2 \approx 1.0233$

- **Obtenção da estimativa de MV de σ**

Passo 1 — Função de verosimilhança

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

Passo 2 — Função de log-verosimilhança

$$\ln L(\sigma | \underline{x}) = -n \ln(\sigma) - \frac{n}{2} \ln(2\pi) - \sum_{i=1}^n \ln(x_i) - \frac{1}{2\sigma^2} \sum_{i=1}^n [\ln(x_i) - 2]^2$$

Passo 3 — Maximização

A estimativa de MV de σ será representada por $\hat{\sigma}$ e

$$\hat{\sigma} : \begin{cases} \frac{d \ln L(\sigma | \underline{x})}{d\sigma} \Big|_{\sigma=\hat{\sigma}} = 0 & \text{(ponto de estacionaridade)} \\ \frac{d^2 \ln L(\sigma | \underline{x})}{d\sigma^2} \Big|_{\sigma=\hat{\sigma}} < 0 & \text{(ponto de máximo)} \end{cases} \Leftrightarrow \begin{cases} -\frac{n}{\hat{\sigma}} + \frac{1}{\hat{\sigma}^3} \sum_{i=1}^n [\ln(x_i) - 2]^2 = 0 \\ \frac{n}{\hat{\sigma}^2} - \frac{3}{\hat{\sigma}^4} \sum_{i=1}^n [\ln(x_i) - 2]^2 < 0 \end{cases}$$

$$\hat{\sigma} : \begin{cases} \hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^n [\ln(x_i) - 2]^2} \\ -\frac{2n^2}{\sum_{i=1}^n [\ln(x_i) - 2]^2} < 0 \quad (\text{prop. verdadeira } \forall (x_1, \dots, x_n)) \end{cases}$$

Passo 4 — Estimativa de MV de β

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^n [\ln(x_i) - 2]^2} \simeq \sqrt{\frac{1}{100} \times 1.0233} = \sqrt{0.010233} \simeq 0.101158.$$

- **Outro parâmetro**

$$h(\sigma) = E(X) = \exp(2 + \sigma^2/2)$$

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

Pela propriedade de invariância dos EMV, temos

$$\widehat{h(\sigma)} = h(\hat{\sigma}) = \exp\left(2 + \frac{\hat{\sigma}^2}{2}\right) \simeq \exp\left(2 + \frac{\sqrt{0.010233}^2}{2}\right) \simeq 7.43.$$

Pergunta 7

2 valores

Assuma que o tempo de autonomia das baterias produzidas por determinada empresa possui distribuição exponencial com valor esperado desconhecido μ . De modo a estimar μ , decidiu-se colocar 64 baterias em teste, tendo-se obtido um tempo médio de autonomia de 29.75 horas.

Recorra à variável aleatória fulcral para μ , $\sqrt{n}[(\bar{X}/\mu) - 1] \stackrel{a}{\sim} \text{normal}(0, 1)$, para construir um intervalo de confiança aproximado a 94% para μ .

- **V.a. de interesse**

X = tempo de autonomia de bateria

- **Situação**

$X \sim \text{exponencial}\left(\frac{1}{\mu}\right)$, pois $E(X) = \mu$

μ DESCONHECIDO

$n = 64 \geq 30$

- **Obtenção de IC aproximado para μ**

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

$$Z = \sqrt{n} \left(\frac{\bar{X}}{\mu} - 1 \right) \stackrel{a}{\sim} \text{normal}(0, 1)$$

Passo 2 — Obtenção dos quantis de probabilidade

Dado que $(1 - \alpha) \times 100\% = 94\%$ (i. e., $\alpha = 0.06$), faremos uso dos quantis simétricos

$$\begin{cases} a_\alpha = -b_\alpha \\ b_\alpha = \Phi^{-1}(1 - \alpha/2) = \Phi^{-1}(0.97) \stackrel{\text{tabela/calc.}}{=} 1.8808. \end{cases}$$

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

$$P \left[a_\alpha \leq Z = \sqrt{n} \left(\frac{\bar{X}}{\mu} - 1 \right) \leq b_\alpha \right] \simeq 1 - \alpha \quad \Leftrightarrow \quad P \left[-\frac{b_\alpha}{\sqrt{n}} + 1 \leq \frac{\bar{X}}{\mu} \leq \frac{b_\alpha}{\sqrt{n}} + 1 \right] \simeq 1 - \alpha$$

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

Passo 4 — IC aproximado a $(1 - \alpha) \times 100\%$ para μ

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

$$IC_{(1-\alpha) \times 100\%}(\mu) \simeq \left[\frac{\bar{x}}{1 + \frac{b_\alpha}{\sqrt{n}}}, \frac{\bar{x}}{1 - \frac{b_\alpha}{\sqrt{n}}} \right].$$

Dado que $n = 64$, $b_\alpha = \Phi^{-1}(1 - \alpha/2) = 1.8808$ e $\bar{x} = 29.75$, temos

$$IC_{94\%}(\mu) \simeq \left[\frac{29.75}{1 + \frac{1.8808}{\sqrt{64}}}, \frac{29.75}{1 - \frac{1.8808}{\sqrt{64}}} \right] \simeq [24.087, 38.894].$$

Pergunta 8

2 valores

Considere que a variável aleatória X representa o erro associado à medição da velocidade de determinado veículo aeroespacial. Suponha que foram anotados os erros associados a 9 medições casuais da velocidade de tal veículo, tendo-se obtido uma média e uma variância (corrigida) amostrais iguais a 11.1 e 0.74 (respectivamente).

Após ter admitido que X possui distribuição normal com valor esperado μ e desvio padrão σ desconhecidos, confronte as hipóteses $H_0: \mu = 10$ e $H_1: \mu \neq 10$ ao nível de significância de 5%.

- V.a. de interesse**

X = erro associado à medição da velocidade de determinado veículo aeroespacial

- Situação**

$X \sim \text{normal}(\mu, \sigma^2)$

μ desconhecido

σ DESCONHECIDO

- Hipóteses**

$H_0: \mu = \mu_0 = 10$

$H_1: \mu \neq \mu_0$

- N.s.**

$\alpha_0 = 5\%$

- Estatística de teste**

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \underset{H_0}{\sim} t_{n-1}$$

- Região de rejeição de H_0 (para valores da estatística de teste)**

O teste é bilateral ($H_1: \mu \neq \mu_0$), logo a região de rejeição de H_0 é do tipo $W = (-\infty, -c) \cup (c, +\infty)$, onde

$$c = F_{t_{(n-1)}}^{-1}(1 - \alpha_0/2) = F_{t_{(8)}}^{-1}(1 - 0.05/2) = F_{t_{(8)}}^{-1}(0.975) \stackrel{\text{tabela/calcul.}}{=} 2.306$$

[para que a prob. de cometer erro de 1a. espécie seja igual a α_0 , i.e., $P(\text{Rejeitar } H_0 \mid H_0) = \alpha_0$].

- Decisão**

Dado que $n = 9$, $\bar{x} = 11.1$ e $s^2 = 0.74$, o valor observado da estatística de teste é

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{11.1 - 10}{\sqrt{0.74/9}} \simeq 3.8362.$$

Uma vez que $t = 3.8362 \in W = (-\infty, -2.306) \cup (2.306, +\infty)$ devemos rejeitar H_0 ao n.s. $\alpha_0 = 5\%$ [ou a qualquer n.s. superior a α_0].

Pergunta 9

2 valores

Um engenheiro conjectura que a velocidade do vento (X , em m/s) sobre uma dada estrutura de engenharia possui função de densidade de probabilidade dada por $f_0(x) = (x/25)e^{-x^2/50}$, para $x \geq 0$ (hipótese H_0). O engenheiro realizou, de forma independente, 100 medições de X e classificou-as de acordo com a tabela apresentada a seguir, tendo incluído também as frequências absolutas esperadas sob H_0 (arredondadas à décima) para cada uma das 6 classes em que foi dividido o contradomínio de X .

Classe	[0, 2[	[2, 4[	[4, 6[	[6, 8[	[8, 10[	[10, +∞[
Freq. absoluta observada	9	12	20	27	18	14
Freq. absoluta esperada sob H_0	7.7	19.7	23.9	20.9	14.3	13.5

Relativamente ao teste de ajustamento de qui-quadrado, para quais dos níveis usuais de significância, 1%, 5% e 10%, deve o engenheiro rejeitar H_0 ?

A: Todos B: 5% e 10% C: 10% D: Nenhum

- V.a. de interesse**

X = velocidade do vento (em m/s) sobre a estrutura de engenharia

- Hipóteses**

$$H_0 : f_X(x) = f_0(x), \quad \forall x$$

$$H_1 : f_X(x) \neq f_0(x), \quad \text{para algum } x$$

- 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 = 6; 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)

Ao efectuarmos um teste de ajustamento do qui-quadrado, a região de rejeição de H_0 é um intervalo à direita $W = (c, +\infty)$. E, ao considerarmos o nível usual de significância α_0 , temos

$$c = F_{\chi^2_{(k-1)}}^{-1}(1 - \alpha_0) = F_{\chi^2_{(5)}}^{-1}(1 - \alpha_0) = \begin{cases} 20.51, & \alpha = 1\% \\ 11.07, & \alpha = 5\% \\ 9.236, & \alpha = 10\% \end{cases}$$

- Decisão** (com base no valor observado da estatística de teste)

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, 2[	9	7.7	$\frac{(9-7.7)^2}{7.7} \approx 0.219$
2	[2, 4[	12	19.7	3.010
3	[4, 6[	20	23.9	0.636
4	[6, 8[	27	20.9	1.780
5	[8, 10[	18	14.3	0.957
6	[10, ∞[	14	13.5	0.019
		$\sum_{i=1}^k o_i = n = 100$	$\sum_{i=1}^k E_i = n = 100$	$t = \sum_{i=1}^k \frac{(o_i - E_i)^2}{E_i} \approx 6.621$

Como $t \simeq 6.6217 < F_{\chi^2_{(5)}}^{-1}(1 - 0.1) = 9.236$, não se rejeita H_0 ao nível usual de significância (n.u.s.) de 10%, nem a qualquer n.s. inferior a 10%, nomeadamente aos n.u.s. de 5% e 1%. Deste modo, a resposta correcta é **D: Nenhum**.

[Em alternativa, poderíamos tomar uma decisão com base no valor-p aproximado: $\text{valor} - p = P(T > t | H_0) \simeq 1 - F_{\chi^2_{(6-1)}}(6.621) \stackrel{\text{calc.}}{\simeq} 0.2504$. Logo, devemos decidir pela: rejeição de H_0 a qualquer n.s. $\alpha_0 > 25.04\%$; não rejeição de H_0 a qualquer n.s. $\alpha_0 \leq 25.04\%$, nomeadamente ao n.u.s. de 1%, 5% e 10%. Por consequência, a resposta, das quatro apresentadas, que se coaduna o referido acima é **D: Nenhum**.

Escusado será dizer que a consulta das tabelas dos quantis de probabilidade da distribuição do qui-quadrado com $k - 1 = 5$ graus de liberdade permite chegar à mesma conclusão. Como $F_{\chi^2_{(6-1)}}^{-1}(0.70) = 6.064 < 6.621 < 7.289 = F_{\chi^2_{(6-1)}}^{-1}(0.80) \Leftrightarrow 0.20 = 1 - 0.80 < 1 - F_{\chi^2_{(6-1)}}(6.621) < 1 - 0.70 = 0.30$, devemos: rejeitar H_0 a qualquer n.s. $\alpha_0 \geq 30\%$; não rejeitar de H_0 a qualquer n.s. $\alpha_0 \leq 20\%$, nomeadamente aos n.u.s. de 1%, 5% e 10%. Assim sendo, a resposta correcta é **D: Nenhum**.]

Pergunta 10

2 valores

Admita que o modelo de regressão linear simples $Y = \beta_0 + \beta_1 x + \epsilon$ foi usado por uma engenheira mecânica para explicar a profundidade do desgaste (Y , em mm) de um tipo de peça mecânica em função do tempo de uso (x , em horas). Para tal, foram recolhidas observações em 25 peças com diferentes tempos de uso, tendo-se obtido os seguintes valores:

$$\sum_{i=1}^{25} x_i = 750, \quad \sum_{i=1}^{25} x_i^2 = 27\,500, \quad \sum_{i=1}^{25} y_i = 90, \quad \sum_{i=1}^{25} y_i^2 = 360, \quad \sum_{i=1}^{25} x_i y_i = 3\,100.$$

Tendo em conta estes valores e as hipóteses de trabalho convenientes, obtenha um intervalo de confiança a 90% para β_1 .

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

Y = profundidade do desgaste (em mm) (v.a. resposta)

x = tempo de uso (em horas) (variável explicativa)

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

$$\epsilon_i \stackrel{i.i.d.}{\sim} \text{normal}(0, \sigma^2), \quad i = 1, \dots, n$$

- **Estimativa de β_1**

Temos $n = 25$ e

- $\sum_{i=1}^n x_i = 750$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{750}{25} = 30$$

$$\sum_{i=1}^n x_i^2 = 27\,500$$

$$\sum_{i=1}^n x_i^2 - n \bar{x}^2 = 27\,500 - 25 \times 30^2 = 5\,000$$

- $\sum_{i=1}^n y_i = 90$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{90}{25} = 3.6$$

$$\sum_{i=1}^n y_i^2 = 360$$

$$\sum_{i=1}^n y_i^2 - n \bar{y}^2 = 360 - 25 \times 3.6^2 = 36$$

- $\sum_{i=1}^n x_i y_i = 3\,100$

$$\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y} = 3\,100 - 25 \times 30 \times 3.6 = 400.$$

Logo, as estimativas de β_1 e da variância dos erros são dadas por:

$$\begin{aligned}\hat{\beta}_1 &= \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sum_{i=1}^n x_i^2 - n \bar{x}^2} = \frac{400}{5000} = 0.08; \\ \hat{\sigma}^2 &= \frac{1}{n-2} [(\sum y_i^2 - n \bar{y}^2) - \hat{\beta}_1^2 (\sum x_i^2 - n \bar{x}^2)] \\ &= \frac{1}{25-2} (36 - 0.08^2 \times 5000) = \frac{4}{23} \simeq 0.173913.\end{aligned}$$

• **Obtenção do IC a 90% para β_1**

Passo 1 — V.a. fulcral para β_1

$$Z = \frac{\hat{\beta}_1 - \beta_1}{se(\hat{\beta}_1)} \sim t_{(n-2)}, \quad \text{onde} \quad se(\hat{\beta}_1) = \sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n \bar{x}^2}}$$

Passo 2 — Quantis de probabilidade

Já que $(1 - \alpha) \times 100\% = 90\%$, temos $\alpha = 0.1$ e lidaremos com

$$\begin{cases} a_\alpha = -b_\alpha \\ b_\alpha = F_{t_{(n-2)}}^{-1}(1 - \alpha/2) = F_{t_{(25-2)}}^{-1}(1 - 0.1/2) = F_{t_{(23)}}^{-1}(0.95) \stackrel{tabela/calc.}{=} 1.714 \end{cases}$$

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

$$\begin{aligned}P(-b_\alpha \leq Z \leq b_\alpha) &= 1 - \alpha \\ P\left[-b_\alpha \leq \frac{\hat{\beta}_1 - \beta_1}{se(\hat{\beta}_1)} \leq b_\alpha\right] &= 1 - \alpha \\ P[\hat{\beta}_1 - b_\alpha \times se(\hat{\beta}_1) \leq \beta_1 \leq \hat{\beta}_1 + b_\alpha \times se(\hat{\beta}_1)] &= 1 - \alpha\end{aligned}$$

Passo 4 — Concretização

A expressão geral do $IC_{(1-\alpha) \times 100\%}(\beta_1)$ é dada por $\left[\hat{\beta}_1 \pm b_\alpha \times \sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n \bar{x}^2}}\right]$ e $\hat{\sigma}^2 \simeq 0.173913$. Onde,

$$IC_{95\%}(\beta_1) = \left[0.08 \pm 1.714 \times \sqrt{\frac{0.173913}{5000}}\right] \simeq [0.069891, 0.090109].$$