

Duração: 120 minutos

Exame de Época de Recurso — A

- 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

Uma empresa classifica os sistemas de segurança dos seus servidores em três categorias distintas: C_1 (alto risco); C_2 (risco médio); C_3 (baixo risco). De acordo com registos recentes: 20% dos servidores pertencem à categoria C_1 e 40% à categoria C_2 ; 30% dos servidores da categoria C_1 já sofreram ataques cibernéticos; 90% dos servidores da categoria C_2 nunca sofreram ataques cibernéticos; e apenas 1% dos servidores da categoria C_3 já sofreram ataques cibernéticos.

Escolhido ao acaso um dos servidores dessa empresa, qual é a probabilidade de ele pertencer à categoria C_2 ou à categoria C_3 , sabendo que esse servidor nunca sofreu ataques cibernéticos?

• **Quadro de acontecimentos e probabilidades**

Acontecimento	Probabilidade
C_1 = servidor pertencer à categoria C_1	$P(C_1) = 0.2$
C_2 = servidor pertencer à categoria C_2	$P(C_2) = 0.4$
C_3 = servidor pertencer à categoria C_3	$P(C_3) = 1 - P(C_1) - P(C_2) = 0.4$
A = servidor ter sofrido ataques cibernéticos	$P(A) = ?$
	$P(A C_1) = 0.3$
	$P(\bar{A} C_2) = 0.9$
	$P(A C_3) = 0.01$

• **Prob. pedida**

Ao invocarmos a lei da probabilidade total, temos

$$\begin{aligned}
 P(A) &= P(A | C_1) \times P(C_1) + P(A | C_2) \times P(C_2) + P(A | C_3) \times P(C_3) \\
 &= 0.3 \times 0.2 + (1 - 0.9) \times 0.4 + 0.01 \times 0.4 \\
 &= 0.104.
 \end{aligned}$$

Ao notarmos que C_2 e C_3 são eventos disjuntos e invocarmos o teorema de Bayes (TB), temos

$$\begin{aligned}
 P[(C_2 \cup C_3) | \bar{A}] &= P(C_2 | \bar{A}) + P(C_3 | \bar{A}) \\
 &\stackrel{TB}{=} \frac{P(\bar{A} | C_2) \times P(C_2)}{P(\bar{A})} + \frac{P(\bar{A} | C_3) \times P(C_3)}{P(\bar{A})} \\
 &= \frac{P(\bar{A} | C_2) \times P(C_2)}{1 - P(A)} + \frac{[1 - P(A | C_3)] \times P(C_3)}{1 - P(A)} \\
 &= \frac{0.9 \times 0.4}{1 - 0.104} + \frac{(1 - 0.01) \times 0.4}{1 - 0.104} = \frac{0.756}{0.896} \\
 &= 0.84375.
 \end{aligned}$$

Pergunta 2

2 valores

Um sensor numa linha de produção deteta defeitos automaticamente, registando o número de defeitos por hora (X). Considere que $X \sim \text{Poisson}(\lambda)$ e que os defeitos têm impacto na qualidade da produção, com um custo associado de $g(X) = 2X^2 + 3X + 5$ (em unidades convenientes).

Sabendo que $\ln[P(X=0)] = -2$, o valor de $E[g(X)]$ é

A: 19 B: 21 C: 23 D: 25

- **V.a.**

X = número de defeitos por hora

- **Distribuição, f.p. e momentos**

$X \sim \text{Poisson}(\lambda)$

$$P(X=x) = \frac{e^{-\lambda} \lambda^x}{x!}, x=0,1,2,\dots$$

$$E(X) = \lambda, V(X) = \lambda$$

- **Cálculo de λ**

$$\lambda : \ln[P(X=0)] = -2 \Leftrightarrow \ln(e^{-\lambda}) = -2 \Leftrightarrow \lambda = 2.$$

- **Valor esperado pedido**

Pela linearidade do valor esperado, vem

$$\begin{aligned} E[g(X)] &= E(2X^2 + 3X + 5) \\ &= 2E(X^2) + 3E(X) + 5 \\ &= 2 \times [V(X) + E^2(X)] + 3E(X) + 5 \\ &\approx 2 \times (2 + 2^2) + 3 \times 2 + 5 \\ &= 23. \end{aligned}$$

Assim sendo, a resposta correcta é **C: 23**.

Pergunta 3

2 valores

Uma fábrica utiliza motores industriais cujos tempos até avaria (em milhares de horas de operação contínua) seguem uma distribuição exponencial. A probabilidade de um desses motores funcionar mais de 2000 horas sem avariar é igual a e^{-2} .

Sabendo que um motor funcionou 1500 horas sem avariar, qual é a probabilidade de o mesmo avariar no decurso das próximas 1000 horas?

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

X = tempos até avaria de um motor, em milhares de horas

$X \sim \text{exponencial}(\lambda)$, com λ a determinar, sendo $P(X > 2) = e^{-2}$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{c.c.}; \end{cases} \quad F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(t) dt = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \geq 0 \end{cases}$$

- **Determinação de λ**

$$\lambda : P(X > 2) = e^{-2} \Leftrightarrow 1 - (1 - e^{-2\lambda}) = e^{-2} \Leftrightarrow -2\lambda = -2 \Leftrightarrow \lambda = 1$$

- **Probabilidade pedida**

$$P(X < 1.5 + 1 \mid X > 1.5) = P(X - 1.5 < 1 \mid X > 1.5) = P(X < 1) = 1 - e^{-1} \approx 0.632121,$$

onde na segunda igualdade se usou a propriedade de falta de memória da distribuição exponencial. [Alternativamente,

$$P(X < 1.5 + 1 \mid X > 1.5) = \frac{P(X < 1.5 + 1, X > 1.5)}{P(X > 1.5)} = \frac{P(1.5 < X < 1.5 + 1)}{1 - P(X \leq 1.5)} = \frac{F_X(1.5 + 1) - F_X(1.5)}{1 - F_X(1.5)} = \frac{(1 - e^{-1.5-1}) - (1 - e^{-1.5})}{1 - (1 - e^{-1.5})} = \frac{e^{-1.5}(1 - e^{-1})}{e^{-1.5}} = 1 - e^{-1} = P(X < 1).]$$

Pergunta 4

2 valores

Considere o par aleatório (X, Y) com função de probabilidade conjunta dada por

$$P(X = x, Y = y) = \begin{cases} xy + \frac{1}{16}, & x = 0, \frac{1}{2}, y = \frac{1}{2}, 1 \\ 0, & \text{outros valores de } x, y. \end{cases}$$

Indique o valor de $E(2X + 1 \mid Y = \frac{1}{2})$.

A: $\frac{5}{6}$ B: $\frac{8}{3}$ C: $\frac{5}{12}$ D: $\frac{11}{6}$

• Par aleatório e f.p. marginais

(X, Y)

Atendendo a que $P(X = x) = \sum_y P(X = x, Y = y)$ e $P(Y = y) = \sum_x P(X = x, Y = y)$, temos

X	Y		P(X = x)
	1/2	1	
0	1/16	1/16	1/8
1/2	5/16	9/16	7/8
P(Y = y)	3/8	5/8	1

• F.p. de X condicional a $Y = 1/2$

$$P(X = x \mid Y = 1/2) = \frac{P(X=x, Y=1/2)}{P(Y=1/2)} = \begin{cases} \frac{P(X=0, Y=1/2)}{P(Y=1/2)} = \frac{1/16}{3/8} = 1/6, & x = 0 \\ \frac{P(X=1/2, Y=1/2)}{P(Y=1/2)} = \frac{5/16}{3/8} = 5/6, & x = 1/2 \\ 0, & \text{outros valores de } x \end{cases}$$

• Valor esperado condicional pedido

Como

$$E(X \mid Y = 1/2) = \sum_{x \in \{0, 1/2\}} x \times P(X = x \mid Y = 1/2) = 0 \times 1/6 + 1/2 \times 5/6 = 5/12,$$

segue-se

$$E(2X + 1 \mid Y = 1/2) = 2 \times E(X \mid Y = 1/2) + 1 = 2 \times 5/12 + 1 = 11/6 = 1.8(3).$$

Donde a resposta corrente seja: **D: $\frac{11}{6}$.**

Pergunta 5

2 valores

Considere que: X_i representa o primeiro algarismo do i -ésimo registo fiscal não fraudulento ($i = 1, 2, \dots, 100$); X_i são variáveis aleatórias independentes e identicamente distribuídas a X ; $E(X) = 3.44$ e $E(X^2) = 17.8917$.

Calcule o valor aproximado de $P(\bar{X} < 3 \text{ ou } \bar{X} > 5)$, onde $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$.

• V.a.; valor esperado e variância comuns

X_i = primeiro dígito do registo fiscal i (em base decimal), $i = 1, \dots, n$

$n = 100$

$X_i \stackrel{\text{indep}}{\sim} X$

$$E(X_i) = \mu \simeq 3.44$$

$$V(X_i) = V(X) = \sigma^2 = E(X^2) - E^2(X) \simeq 17.8917 - (3.44)^2 = 6.0581$$

- **V.a. de interesse**

$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ = média do primeiro dígito em n registos fiscais

- **Valor esperado e variância de $\bar{X}$**

$$E(\bar{X}) = E\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n} \sum_{i=1}^n E(X_i) \stackrel{X_i \sim X}{=} \frac{1}{n} n E(X) = \mu$$

$$V(\bar{X}) = V\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \stackrel{X_i \text{ indep.}}{=} \frac{1}{n^2} \sum_{i=1}^n V(X_i) \stackrel{X_i \sim X}{=} \frac{1}{n^2} n V(X) = \frac{\sigma^2}{n}$$

- **Distribuição aproximada de $\bar{X}$**

De acordo com o teorema do limite central (TLC),

$$\frac{\bar{X} - E(\bar{X})}{\sqrt{V(\bar{X})}} = \frac{\bar{X} - E(X)}{\sqrt{V(X)/n}} = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \stackrel{a}{\sim} \text{normal}(0, 1).$$

- **Prob. pedida** (valor aproximado)

$$\begin{aligned} P(\bar{X} < 3 \text{ ou } \bar{X} > 5) &= P\left(\frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \leq \frac{3 - \mu}{\sqrt{\sigma^2/n}}\right) + 1 - P\left(\frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} \leq \frac{5 - \mu}{\sqrt{\sigma^2/n}}\right) \\ &\stackrel{TLC}{\simeq} \Phi\left(\frac{3 - \mu}{\sqrt{\sigma^2/n}}\right) + 1 - \Phi\left(\frac{5 - \mu}{\sqrt{\sigma^2/n}}\right) \\ &= \Phi\left(\frac{3 - 3.44}{\sqrt{6.0581/100}}\right) + 1 - \Phi\left(\frac{5 - 3.44}{\sqrt{6.0581/100}}\right) \\ &\simeq \Phi(-1.79) + 1 - \Phi(6.34) \\ &= 1 - \Phi(1.79) + 1 - \Phi(6.34) \\ &\stackrel{\text{tabelas/calc.}}{\simeq} 1 - 0.9633 + 1 - 1.0000 \\ &= 0.0367. \end{aligned}$$

Pergunta 6

2 valores

Admita que o tempo, em horas, necessário para reparar uma bomba de calor é representado pela variável aleatória X com função de densidade de probabilidade dada por

$$f_X(x) = \frac{\alpha^3 x^2 e^{-\alpha x}}{2}, \quad x > 0,$$

onde α é um parâmetro a estimar.

Deduza o estimador de máxima verosimilhança de α com base numa amostra aleatória de X de dimensão n , $(X_1, \dots, X_n)$.

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

X = tempo necessário para reparar uma bomba de calor (horas)

$$f_X(x) = \frac{\alpha^3 x^2 e^{-\alpha x}}{2}, \quad x > 0$$

- **Parâmetro desconhecido**

$$\alpha \quad (\alpha > 0)$$

- **Amostra; amostra aleatória**

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

$(X_1, X_2, \dots, X_n)$, amostra aleatória de dimensão n proveniente de X .

- **Dedução do estimador de MV de α**

Passo 1 — Função de verosimilhança

$$L(\alpha | \underline{x}) = f_{\underline{X}}(\underline{x}) \stackrel{X_i \text{ indep}}{=} \prod_{i=1}^n f_{X_i}(x_i) \stackrel{X_i \sim X}{=} \prod_{i=1}^n f_X(x_i) = \prod_{i=1}^n \left(\frac{\alpha^3 x_i^2 e^{-\alpha x_i}}{2} \right) = 2^{-n} \alpha^{3n} e^{-\alpha \sum_{i=1}^n x_i} \prod_{i=1}^n x_i^2, \quad \alpha > 0$$

Passo 2 — Função de log-verosimilhança

$$\ln L(\alpha | \underline{x}) = -n \ln(2) + 3n \ln(\alpha) - \alpha \sum_{i=1}^n x_i + 2 \sum_{i=1}^n \ln(x_i), \quad \alpha > 0$$

Passo 3 — Maximização

A estimativa de MV de α será representada por $\hat{\alpha}$ e

$$\hat{\alpha} : \begin{cases} \left. \frac{d \ln L(\alpha | \underline{x})}{d\alpha} \right|_{\alpha=\hat{\alpha}} = 0 & \text{(ponto de estacionaridade)} \\ \left. \frac{d^2 \ln L(\alpha | \underline{x})}{d\alpha^2} \right|_{\alpha=\hat{\alpha}} < 0 & \text{(ponto de máximo)} \end{cases}$$

$$\begin{cases} \frac{3n}{\hat{\alpha}} - \sum_{i=1}^n x_i = 0 \\ -\frac{3n}{\hat{\alpha}^2} < 0 \end{cases} \Leftrightarrow \begin{cases} \hat{\alpha} = \frac{3n}{\sum_{i=1}^n x_i} \\ -\frac{3n}{\hat{\alpha}^2} < 0 \quad (\text{prop. verdadeira}) \end{cases}$$

Passo 4 — Estimador de MV de α

$$EMV(\alpha) = \frac{3n}{\sum_{i=1}^n X_i} = \frac{3}{\bar{X}}.$$

Pergunta 7

2 valores

O método usado na medição do pH certo processo químico fornece valores que se admite terem distribuição normal de valor esperado μ e desvio padrão σ desconhecidos. Para estimar μ , efectuaram-se 10 medições independentes tendo-se obtido um total amostral igual a 78.08 e um desvio padrão amostral $s = 0.498$.

Com base na amostra recolhida, o intervalo de confiança a 90% para μ é

A: [7.52, 8.10] **B:** [7.59, 8.03] **C:** [7.55, 8.07] **D:** [7.40, 8.22]

- **V.a. de interesse**

X = valor da medição do pH de uma solução reagente

- **Situação**

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

$\mu = E(X)$ DESCONHECIDO

$\sigma^2 = V(X)$ desconhecida

- **Obtenção de IC aproximado para μ**

Passo 1 - Seleção da v.a. fulcral

$$Z = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{(n-1)}$$

Passo 2 — Obtenção dos quantis de probabilidade

Como $(1 - \alpha) \times 100\% = 90\% \Leftrightarrow \alpha = 0.1$, lidaremos com os quantis seguintes:

$$\begin{cases} a_{\alpha} = F_{t_{(n-1)}}^{-1}(\alpha/2) = -F_{t_{(n-1)}}^{-1}(1 - \alpha/2) = -F_{t_{(9)}}^{-1}(0.95) \stackrel{\text{tabela/calcul.}}{=} -1.833 \\ b_{\alpha} = F_{t_{(n-1)}}^{-1}(1 - \alpha/2) = F_{t_{(9)}}^{-1}(0.95) = 1.833. \end{cases}$$

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

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

$$P\left(a_\alpha \leq \frac{\bar{X} - \mu}{S/\sqrt{n}} \leq b_\alpha\right) = 1 - \alpha$$

$$P\left(\bar{X} - b_\alpha \times \frac{S}{\sqrt{n}} \leq \mu \leq \bar{X} - a_\alpha \times \frac{S}{\sqrt{n}}\right) = 1 - \alpha$$

Passo 4 — Concretização

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

$$IC_{(1-\alpha) \times 100\%}(\mu) = \left[\bar{x} - F_{t(n-1)}^{-1}(1 - \alpha/2) \times \frac{s}{\sqrt{n}}, \bar{x} + F_{t(n-1)}^{-1}(1 - \alpha/2) \times \frac{s}{\sqrt{n}} \right].$$

Atendendo a que $n = 10$, $\alpha = 0.1$, $F_{t(n-1)}^{-1}(1 - \alpha/2) = 1.833$, $\bar{x} = \frac{78.08}{10} = 7.808$ e $s = 0.498$, segue-se

$$IC_{95\%}(\mu) = \left[7.808 - 1.833 \times \frac{0.498}{\sqrt{10}}, 7.808 + 1.833 \times \frac{0.498}{\sqrt{10}} \right] \simeq [7.52, 8.10].$$

Logo, a resposta correcta é **A**: [7.52, 8.10].

Pergunta 8

2 valores

Uma fábrica de componentes eletrónicas produz resistores. Após efetuar modificações no processo de produção, um engenheiro decide averiguar se o desvio padrão da resistência eléctrica (**normalmente distribuída**) dos resistores é inferior a 4Ω , tendo selecionado casualmente e ensaiado 30 desses resistores e obtido uma variância amostral corrigida de $s^2 = 10.24 \Omega^2$.

Confronte as hipóteses $H_0 : \sigma = 4$ e $H_1 : \sigma < 4$, obtendo para o efeito o valor-p (ou um intervalo para o valor-p) do teste. Para que níveis de significância deve o engenheiro rejeitar H_0 ?

- V.a. de interesse**

X = resistência do resistor

- Situação**

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

μ desconhecido

σ DESCONHECIDO

- Hipóteses**

$H_0 : \sigma = \sigma_0 = 4$ vs. $H_1 : \sigma < \sigma_0$

- Estatística de teste**

$$\frac{(n-1)S^2}{\sigma_0^2} \sim_{H_0} \chi_{(n-1)}^2$$

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

Uma vez que o teste é unilateral inferior ($H_1 : \sigma < \sigma_0$), a região de rejeição de H_0 é do tipo $W = (0, c)$.

- Decisão (com base no valor-p)**

Como o valor observado da estatística de teste é igual a

$$t = \frac{(n-1)s^2}{\sigma_0^2} = \frac{(30-1) \times 10.24}{4^2} \simeq 18.56.$$

e $W = (0, c)$, o valor-p é dado por

$$valor - p = P(T < t | H_0) = F_{\chi^2_{(n-1)}}(t) = F_{\chi^2_{(29)}}(18.56) \stackrel{calc.}{=} 0.067949.$$

Logo, o engenheiro deve:

- [- não rejeitar H_0 a qualquer n.s. $\alpha_0 \leq valor - p = 6.7949\%$, em particular aos n.u.s. de 1% e 5%;]
- rejeitar H_0 , a qualquer n.s. $\alpha_0 > valor - p = 6.7949\%$, designadamente ao n.u.s. de 10%.

[Assim sendo, a conjectura do engenheiro (H_1) só é consistente com os dados ao n.u.s. de 10%.]
[Alternativamente, poderíamos recorrer às tabelas de quantis da distribuição do qui-quadrado e adiantar um intervalo para o valor-p:

$$F_{\chi^2_{(29)}}^{-1}(0.05) = 17.71 < t = 18.56 < 18.85 = F_{\chi^2_{(29)}}^{-1}(0.075)$$

$$0.05 < valor - p = F_{\chi^2_{(29)}}(18.56) < 0.075$$

Assim, é suposto: [não rejeitar H_0 a qualquer n.s. $\alpha_0 \leq valor - p = 5\%$, designadamente aos n.u.s. de 1% e 5%;] rejeitar H_0 a qualquer n.s. $\alpha_0 \geq valor - p = 7.5\%$, nomeadamente ao n.u.s. de 10%.]

Pergunta 9

2 valores

Seja X a variável aleatória que representa o número de ensaios, em 4 realizados, que resultam na separação correta de determinadas partículas. Uma engenheira química conjectura que X possui distribuição binomial com parâmetros 4 e 0.5 (hipótese H_0). Tendo recolhido uma amostra casual de 100 valores de X , a engenheira organizou uma tabela com as seguintes frequências absolutas observadas e esperadas sob H_0 :

X	0	1	2	3	4
Frequência absoluta observada	2	6	32	33	27
Frequência absoluta esperada sob H_0	6.25	25	37.5	25	6.25

Efetue o teste de ajustamento do qui-quadrado, considerando para o efeito o nível de significância de 2.5%.

- **V.a. de interesse**

X = número de ensaios, em 4, que resultam na separação correta de determinadas partículas

- **Hipóteses**

$H_0 : X \sim \text{binomial}(4, 0.5)$

$H_1 : X \not\sim \text{binomial}(4, 0.5)$

- **N.s.**

$\alpha_0 = 2.5\%$

- **Estatística de teste**

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

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 observados de T)

Ao efectuarmos um teste de ajustamento do qui-quadrado, a região de rejeição de H_0 é um intervalo à direita $W = (c, +\infty)$, onde

$$c = F_{\chi^2_{(k-1)}}^{-1}(1 - \alpha_0) = F_{\chi^2_{(5-1)}}^{-1}(1 - 0.025) = F_{\chi^2_{(4)}}^{-1}(0.975) = 11.14.$$

• **Decisão**

i	Classe	Freq. abs. obs. o_i	Freq. abs. esper. sob H_0 E_i	Parcelas valor obs. estat. teste $\frac{(o_i - E_i)^2}{E_i}$
1	0	2	6.25	$\frac{(2-6.25)^2}{6.25} \approx 2.89$
2	1	6	25	14.44
3	2	32	37.5	0.81
4	3	33	25	2.56
5	4	27	6.25	68.89
		$\sum_{i=1}^k o_i = n = 100$	$\sum_{i=1}^k E_i = n = 100$	$t = \sum_{i=1}^k \frac{(o_i - E_i)^2}{E_i} = 89.59$

Uma vez que $t = 89.59 \in W = (11.14, +\infty)$ devemos rejeitar H_0 ao n.s. $\alpha_0 = 2.5\%$ [ou a qualquer n.s. superior a α_0].

Pergunta 10

2 valores

Após se ter admitido a validade do modelo de regressão linear simples, $Y = \beta_0 + \beta_1 x + \epsilon$, foram obtidas as seguintes quantidades para um conjunto de 20 pares de observações do teor global de humidade de uma mistura (x , em percentagem) e a densidade do produto final (Y , em g/cm³):

$$\sum_{i=1}^{20} x_i^2 - 20\bar{x}^2 = 170, \quad \hat{\beta}_0 = -0.91, \quad \hat{\beta}_1 = 0.74, \quad \hat{\sigma}^2 \approx 0.1267.$$

Tendo presente estes valores e as hipóteses de trabalho convenientes, teste $H_0 : \beta_1 = 0$ contra $H_1 : \beta_1 > 0$. 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%.

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

Y = densidade do produto final (v.a. resposta)

x = teor global de humidade da mistura (variável explicativa)

$$Y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \quad i = 1, \dots, n, \quad (n = 20)$$

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

• **Hipóteses**

$$H_0 : \beta_1 = \beta_{1,0} = 0$$

$$H_1 : \beta_1 > \beta_{1,0}$$

• **Estatística de teste**

$$T = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}}} \sim_{H_0} t_{(n-2)}$$

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

Estamos a lidar com um teste unilateral superior ($H_1 : \beta_1 > \beta_{1,0}$), logo a região de rejeição de H_0 é do tipo $W = (c, +\infty)$.

• **Decisão** (com base no valor-p)

Atendendo ao enunciado, $\sum_{i=1}^{20} x_i^2 - 20\bar{x}^2 = 170$, $\hat{\beta}_1 = 0.74$ e $\hat{\sigma}^2 \approx 0.1267$. Logo, o valor observado da estatística de teste e o valor-p são iguais a

$$t = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}}} \simeq \frac{0.74 - 0}{\sqrt{\frac{0.1267}{170}}} \simeq 27.109719.$$

$$\text{valor-p} = P(T > t \mid H_0) = 1 - F_{t_{(n-2)}}(t) = 1 - F_{t_{(18)}}(27.109719) \stackrel{calc.}{\simeq} 0.$$

Logo, devemos

- rejeitar H_0 a qualquer n.s. $\alpha_0 > \text{valor-p} \simeq 0\%$, designadamente aos n.u.s. de 1%, 5% e 10%.

[Alternativamente, poderíamos recorrer às tabelas de quantis da distribuição do qui-quadrado e adiantar um intervalo para o valor-p:

$$\begin{aligned} F_{t_{(18)}}^{-1}(0.9995) = 9.922 &< t \simeq 27.109719 < +\infty = F_{t_{(18)}}^{-1}(1) \\ 0.9995 &< F_{t_{(18)}}(27.109719) < 1 \\ 0 &< \text{valor-p} = 1 - F_{t_{(18)}}(27.109719) < 0.0005. \end{aligned}$$

Assim sendo, é suposto

- rejeitar H_0 a qualquer n.s. $\alpha_0 \geq 0.05\%$, nomeadamente ao n.u.s. de 1%, 5% e 10% [opção **A**].