

Duração: 120 minutos

Exame Época Recurso – B

- As questões de escolha múltipla têm uma e apenas uma resposta correta. Caso assinale uma resposta errada, serão descontados 0.5 valores.
- Se assinalar mais do que uma resposta numa questão de escolha múltipla, o resultado será classificado como errado e 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.

Pergunta 1

2 valores

Um sistema electrónico com três componentes (1, 2 e 3) funciona se e só se pelo menos duas dessas componentes estiverem operacionais, sendo que a probabilidade de a componente 1 (respetivamente 2 e 3) estar operacional é igual a 0.9 (respetivamente 0.8 e 0.7). Admita que as componentes estão (ou não) operacionais independentemente umas das outras.

Calcule a probabilidade de o sistema estar em funcionamento.

• **Quadro de acontecimentos e probabilidades**

Acontecimento	Probabilidade
A_1 = “a componente 1 está operacional”	$P(A_1) = 0.9$
A_2 = “a componente 2 está operacional”	$P(A_2) = 0.8$
A_3 = “a componente 3 está operacional”	$P(A_3) = 0.7$
F = “o sistema está em funcionamento”	$P(F) = ?$

• **Cálculo da probabilidade pedida**

Tirando partido da independência (mútua) entre os acontecimentos A_1 , A_2 e A_3 e do facto de $(A_1 \cap A_2 \cap A_3)$, $(A_1 \cap A_2 \cap \bar{A}_3)$, $(A_1 \cap \bar{A}_2 \cap A_3)$ e $(\bar{A}_1 \cap A_2 \cap A_3)$ serem acontecimentos disjuntos, tem-se:

$$\begin{aligned} P(F) &= P[(A_1 \cap A_2 \cap A_3) \cup (A_1 \cap A_2 \cap \bar{A}_3) \cup (A_1 \cap \bar{A}_2 \cap A_3) \cup (\bar{A}_1 \cap A_2 \cap A_3)] \\ &= P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap \bar{A}_3) + P(A_1 \cap \bar{A}_2 \cap A_3) + P(\bar{A}_1 \cap A_2 \cap A_3) \\ &= P(A_1) \times P(A_2) \times P(A_3) + P(A_1) \times P(A_2) \times P(\bar{A}_3) + P(A_1) \times P(\bar{A}_2) \times P(A_3) \\ &\quad + P(\bar{A}_1) \times P(A_2) \times P(A_3) \\ &= 0.9 \times 0.8 \times 0.7 + 0.9 \times 0.8 \times (1 - 0.7) + 0.9 \times (1 - 0.8) \times 0.7 + (1 - 0.9) \times 0.8 \times 0.7 \\ &= 0.902. \end{aligned}$$

Pergunta 2

2 valores

O número de jogos de computador vendidos diariamente por uma loja *online* é uma variável aleatória X com distribuição de Poisson com variância igual a 1.

Assumindo que os números de jogos de computador vendidos em dias distintos são variáveis aleatórias independentes e identicamente distribuídas a X , qual é a probabilidade de que em 7 dias a loja venda mais de 10 jogos?

A: 0.8305 B: 0.0985 C: 0.9015 D: 0.1695

- **V.a.**

X_1 = número diário de jogos vendidos pela loja online

$X_1 \sim \text{Poisson}(\lambda)$, onde $\lambda: V(X_1) = 1 \Leftrightarrow \lambda = 1$

- **V.a. de interesse**

X_7 = número de jogos vendidos pela loja online em 7 dias

- **Distribuição**

$X_7 \sim \text{Poisson}(\lambda \times 7 = 7)$, pela propriedade reprodutiva da distribuição de Poisson.

- **F.p. de X**

$$P(X_7 = x) = \frac{e^{-7} \times 7^x}{x!}, \quad x \in \mathbb{N}_0$$

- **Probabilidade pedida**

$$\begin{aligned} P(X_7 > 10) &= 1 - P(X_7 \leq 10) \\ &= 1 - F_{\text{Poisson}(7)}(10) \\ &\stackrel{\text{tabelas, calc.}}{\approx} 1 - 0.9015 \\ &= 0.0985. \end{aligned}$$

Pergunta 3	2 valores
-------------------	-----------

Seja X uma variável aleatória com distribuição uniforme contínua no intervalo $[0, 1]$.

Determine a probabilidade de X ser superior a $\frac{1}{2}$ sabendo que $|X - \frac{1}{2}| < \frac{1}{4}$.

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

$X \sim \text{uniforme}(0, 1)$

$$f_X(x) = \begin{cases} 1, & 0 \leq x \leq 1 \\ 0, & \text{caso contrário.} \end{cases}$$

$$F_X(x) = \int_{-\infty}^x f_X(t) dt = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x \leq 1 \\ 1, & x > 1. \end{cases}$$

- **Probabilidade pedida**

$$\begin{aligned} P\left(X > \frac{1}{2} \mid |X - \frac{1}{2}| < \frac{1}{4}\right) &= \frac{P\left(X > \frac{1}{2}, -\frac{1}{4} < X - \frac{1}{2} < \frac{1}{4}\right)}{P\left(-\frac{1}{4} < X - \frac{1}{2} < \frac{1}{4}\right)} \\ &= \frac{P\left(\frac{1}{2} < X < \frac{3}{4}\right)}{P\left(\frac{1}{4} < X < \frac{3}{4}\right)} = \frac{F_X(3/4) - F_X(1/2)}{F_X(3/4) - F_X(1/4)} = \frac{3/4 - 1/2}{3/4 - 1/4} = \frac{1}{2}. \end{aligned}$$

Pergunta 4	2 valores
-------------------	-----------

Considere que as durações de 50 componentes eletrônicas são variáveis aleatórias independentes com distribuição comum exponencial com valor esperado de 10 dias.

Determine um valor aproximado para o quantil de probabilidade 0.05 da duração total dessas 50 componentes eletrônicas.

- **V.a.; valor esperado e variância comuns**

X_i = duração (em dias) da componente eletrônica i , $i = 1, \dots, n$

$n = 50 > 30$

$X_i \stackrel{i.i.d.}{\sim} X \sim \text{exponencial}(\lambda)$, $i = 1, \dots, n$

$E(X_i) = E(X) = \mu = 1/\lambda = 10$

$V(X_i) = V(X) = \sigma^2 = 1/\lambda^2 = 10^2$

- **V.a. de interesse**

$S_n = \sum_{i=1}^n X_i$ = duração de n componentes eletrônicas

- **Valor esperado e variância de S_n**

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

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

- **Distribuição aproximada de S_n**

Segundo o teorema do limite central (TLC),

$$\frac{S_n - E(S_n)}{\sqrt{V(S_n)}} = \frac{S_n - nE(X)}{\sqrt{nV(X)}} = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \underset{\text{TLC}}{\approx} \text{normal}(0, 1).$$

- **Quantil pedido (valor aproximado)**

$$\xi : P(S_n \leq \xi) = 0.05 \Leftrightarrow P\left(\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \leq \frac{\xi - n\mu}{\sqrt{n\sigma^2}}\right) = 0.05 \Leftrightarrow \Phi\left(\frac{\xi - n\mu}{\sqrt{n\sigma^2}}\right) \stackrel{\text{TLC}}{\approx} 0.05$$

$$\frac{\xi - n\mu}{\sqrt{n\sigma^2}} \approx \Phi^{-1}(0.05) \Leftrightarrow \xi \approx n\mu + \sqrt{n}\sigma \times \Phi^{-1}(0.05)$$

$$\xi \stackrel{\text{tabelas/calc.}}{\approx} 50 \times 10 + \sqrt{50} \times 10 \times (-1.6449)$$

$$\xi \approx 383.688.$$

Pergunta 5	2 valores
-------------------	-----------

Seja (X, Y) um par aleatório discreto com função de probabilidade conjunta

$$P(X = x, Y = y) = \frac{1}{2^{x+y}}, \quad x, y = 1, 2, \dots$$

Determine a função de probabilidade marginal de X .

- **Par aleatório**

(X, Y) com a f.p. conjunta do enunciado

- **F.p. marginal de X**

$$\begin{aligned} P(X = x) &= \sum_y P(X = x, Y = y) \\ &= \sum_{y=1}^{+\infty} \frac{1}{2^{x+y}} = \frac{1}{2^x} \sum_{y=1}^{+\infty} \frac{1}{2^y} = \frac{1}{2^x} \times \frac{1}{2} \frac{1}{1 - \frac{1}{2}} = \frac{1}{2^x}, \quad x \in \mathbb{N}. \end{aligned}$$

Admita que a variável aleatória X , que representa o número de eventos que ocorrem numa região bidimensional com área unitária, possui função de probabilidade

$$P(X = x) = \begin{cases} \frac{e^{-\lambda} \lambda^x}{x!}, & x = 0, 1, 2, \dots \\ 0, & \text{caso contrário,} \end{cases}$$

onde λ é um parâmetro positivo desconhecido.

Deduzo o estimador de máxima verosimilhança de λ , baseado numa amostra aleatória de dimensão n proveniente de X .

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

X

$$P(X = x) = \begin{cases} \frac{e^{-\lambda} \lambda^x}{x!}, & x = 0, 1, 2, \dots \\ 0, & \text{caso contrário,} \end{cases}$$

- **Parâmetro desconhecido**

λ ($\lambda > 0$)

- **Amostra; amostra aleatória**

$\underline{x} = (x_1, \dots, x_n)$, amostra de dimensão n proveniente de X .

$\underline{X} = (X_1, \dots, X_n)$, amostra aleatória de dimensão n proveniente de X .

- **Dedução do estimador de MV de λ**

Passo 1 — Função de verosimilhança

$$\begin{aligned} L(\lambda | \underline{x}) &= P(\underline{X} = \underline{x}) \stackrel{X_i \text{ indep}}{=} \prod_{i=1}^n P(X_i = x_i) \stackrel{X_i \sim X}{=} \prod_{i=1}^n P(X = x_i) \\ &= \prod_{i=1}^n \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} \\ &= e^{-n\lambda} \lambda^{\sum_{i=1}^n x_i} \frac{1}{\prod_{i=1}^n x_i!}, \quad \lambda > 0 \end{aligned}$$

Passo 2 — Função de log-verosimilhança

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

Passo 3 — Maximização

A estimativa de MV de λ é representada por $\hat{\lambda}$ e

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

$$\begin{cases} -n + \frac{\sum_{i=1}^n x_i}{\hat{\lambda}} = 0 \\ -\frac{\sum_{i=1}^n x_i}{\hat{\lambda}^2} < 0 \end{cases} \Leftrightarrow \begin{cases} \hat{\lambda} = \frac{\sum_{i=1}^n x_i}{n} = \bar{x} & \text{(média da amostra)} \\ -\frac{n\bar{x}}{\bar{x}^2} = -\frac{n}{\bar{x}} < 0 & \text{(prop. verdadeira; } \bar{x} > 0) \end{cases}$$

Passo 4 — Estimador de MV de λ

$$EMV(\lambda) = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X} \quad \text{(média da amostra aleatória)}$$

Admita que o peso máximo suportado por uma marquesa portátil ergonómica é uma variável aleatória X com distribuição normal com valor esperado desconhecido e variância igual a 5. Foram selecionadas aleatoriamente 28 dessas marquesas, tendo-se observado uma média amostral de $\bar{x} = 148.5$.

Determine um intervalo de confiança a 95% para o valor esperado de X .

A: [148.805, 150.195] B: [147.672, 149.328] C: [146.648, 150.352] D: [149.946, 153.054]

- **V.a. de interesse**

X = peso máximo suportado por uma marquesa portátil ergonómica

- **Situação**

$X \sim \text{normal}(\mu, \sigma^2)$, $\mu = E(X)$ DESCONHECIDO, $\sigma^2 = V(X) = 5$

- **Obtenção do IC para μ**

Passo 1 — Variável aleatória fulcral para μ

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

Passo 2 — Quantis de probabilidade

Dado que $(1 - \alpha) \times 100\% = 95\%$, usaremos os quantis

$$(a_\alpha, b_\alpha) : \begin{cases} P(a_\alpha \leq Z \leq b_\alpha) = 1 - \alpha \\ P(Z < a_\alpha) = P(Z > b_\alpha) = \alpha/2 \end{cases} \Leftrightarrow \begin{cases} a_\alpha = \Phi^{-1}(\alpha/2) = \Phi^{-1}(0.025) = -1.9600 \\ b_\alpha = \Phi^{-1}(1 - \alpha/2) = \Phi^{-1}(0.975) = 1.9600. \end{cases}$$

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

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

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

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

Passo 4 — Concretização

Atendendo à expressão geral do IC para μ ,

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

ao par de quantis acima e ao facto de $\bar{x} = 148.5$ e $\sigma^2 = 5$, temos:

$$\begin{aligned} IC_{92\%}(\mu) &= \left[148.5 - 1.6449 \times \frac{\sqrt{5}}{\sqrt{28}}, 148.5 + 1.6449 \times \frac{\sqrt{5}}{\sqrt{28}} \right] \\ &\approx [147.672, 149.328]. \end{aligned}$$

Suponha que o comprimento (em mm) de uma peça é uma variável aleatória com distribuição normal com valor esperado μ e desvio padrão σ desconhecidos.

Teste a hipótese $H_0 : \sigma = 1$ contra $H_1 : \sigma < 1$, num dia em que foram recolhidas casualmente 6 peças e a média e a variância corrigida amostrais foram de $\bar{x} = 101.0$ e $s^2 = 0.60$. Decida com base no valor-p.

- **V.a. de interesse**

X = comprimento de uma peça metálica

- **Situação**

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

μ desconhecido

σ DESCONHECIDO

- **Hipóteses**

$H_0 : \sigma = \sigma_0 = 1$ vs. $H_1 : \sigma < \sigma_0$

- **Estatística de teste**

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

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

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

- **Decisão (com base no valor-p)**

Atendendo a que o valor observado da estatística de teste e o valor-p são iguais a

$$\begin{aligned} t &= \frac{(n-1)s^2}{\sigma_0^2} \\ &= \frac{(6-1) \times 0.60}{1^2} \\ &= 3.0 \\ \text{valor-p} &= P(T < t \mid H_0) \\ &= F_{\chi^2_{(n-1)}}(t) \\ &= F_{\chi^2_{(5)}}(3.0) \\ &\stackrel{\text{tabelas, calc.}}{=} 0.30, \end{aligned}$$

é suposto:

- não rejeitar H_0 a qualquer n.s. $\alpha_0 \leq \text{valor-p} = 30\%$, designadamente aos níveis usuais de significância (1%, 5%, 10%);
- rejeitar H_0 a qualquer n.s. $\alpha_0 > \text{valor-p} = 30\%$.

Pergunta 9	2 valores
-------------------	-----------

Após uma fase preliminar de *marketing*, conjectura-se que o número de refeições veganas vendidas diariamente em determinado estabelecimento possui distribuição de Poisson com valor esperado igual a 4 (hipótese H_0). Uma amostra casual conduziu à seguinte tabela de frequências:

Número de refeições veganas vendidas diariamente	{0, 1, 2}	3	4	{5, 6, ...}
Frequência absoluta observada	8	16	10	6
Frequência absoluta esperada sob H_0	9.52	7.81	7.81	14.86

Recorra ao teste de ajustamento do qui-quadrado para testar H_0 ao nível de significância de 5%.

- **V.a. de interesse**

X = número de refeições veganas vendidas diariamente

- **Hipóteses**

$H_0 : X \sim \text{Poisson}(4)$

$H_1 : X \not\sim \text{Poisson}(4)$

- **Nível de significância**

$\alpha_0 = 5\%$

- **Estatística de teste**

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

onde: k = no. de classes = 4; O_i = frequência absoluta observável da classe i ; E_i = frequência absoluta esperada, sob H_0 , da classe i .

- **[Frequências absolutas esperadas sob H_0 — De acordo com a tabela facultada, as frequências absolutas esperadas sob H_0 aproximadas às centésimas são: $E_1 \approx 9.52$; $E_2 \approx 7.81$; $E_3 \approx 7.81$; $E_4 \approx 14.86$. Não é necessário fazer qualquer agrupamento de classes uma vez que se verifica $E_i \geq 5$, em pelo menos 80% das classes, e que $E_i \geq 1$, para todo o i . Caso fosse preciso efectuar agrupamento de classes, os valores de k e $c = F_{\chi^2_{(k-1)}}^{-1}(1 - \alpha_0)$ teriam que ser recalculados...]**

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

Lidamos com um teste de ajustamento, donde a região de rejeição de H_0 é o intervalo à direita $W = (c, +\infty)$, onde

$$c = F_{\chi^2_{(k-\beta-1)}}^{-1}(1 - \alpha_0) = F_{\chi^2_{(4-1)}}^{-1}(1 - 0.05) \stackrel{\text{tabela/calcul.}}{=} 7.815.$$

- **Decisão**

i	Classe i	Freq. abs. obs. o_i	Freq. abs. esp. sob H_0 E_i	Parcelas valor obs. estat. teste $\frac{(o_i - E_i)^2}{E_i}$
1	{0, 1, 2}	8	9.52	$\frac{(8-9.52)^2}{9.52} \approx 0.243$
2	{3}	16	7.81	8.588
3	{4}	10	7.81	0.614
4	{5, 6, ...}	6	14.86	5.283
		$\sum_{i=1}^k o_i = n = 40$	$\sum_{i=1}^k E_i = n = 40$	$t = \sum_{i=1}^k \frac{(o_i - E_i)^2}{E_i} \approx 14.728$

Uma vez que $t \approx 14.728 \in W = (7.815, +\infty)$, devemos rejeitar H_0 ao n.s. de $\alpha_0 = 5\%$ [ou a qualquer outro n.s. superior a α_0].

Pergunta 10

2 valores

O gestor de uma frota de camiões de longo curso registou a distância percorrida (x , em quilómetros) e a quantidade de combustível consumido (Y , em litros por dia). A amostra casual conduziu a:

$$\sum_{i=1}^8 x_i = 2672, \quad \sum_{i=1}^8 x_i^2 = 900356, \quad \sum_{i=1}^8 y_i = 763, \quad \sum_{i=1}^8 y_i^2 = 73395, \quad \sum_{i=1}^8 x_i y_i = 257048, \quad \hat{\sigma}^2 \approx 1.415602.$$

Admita que as variáveis x e Y estão relacionadas de acordo com o modelo de regressão linear simples,

$$Y = \beta_0 + \beta_1 x + \epsilon.$$

Tendo presente as hipóteses de trabalho convenientes, teste a hipótese $H_0 : \beta_1 = 0$ contra $H_1 : \beta_1 \neq 0$, ao nível de significância de 5%.

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

Y = combustível consumido (v.a. resposta)

x = distância percorrida (variável explicativa)

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

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

- **Hipóteses**

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

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

- **Nível de significância**

$$\alpha_0 = 5\%$$

- **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 bilateral ($H_1 : \beta_1 \neq \beta_{1,0}$), logo a região de rejeição de H_0 é do tipo $W = (-\infty, -c) \cup (c, +\infty)$, onde $c : P(\text{Rejeitar } H_0 | H_0) = \alpha_0$, i.e.,

$$c = F_{t_{(n-2)}}^{-1}(1 - \alpha_0/2) = F_{t_{(8-2)}}^{-1}(1 - 0.05/2) \stackrel{\text{tabela/calcul.}}{\approx} 2.447.$$

- **Decisão**

Atendendo a que a estimativa de MQ de β_1 é dada por

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} = \frac{257048 - 8 \times \frac{2672}{8} \times \frac{763}{8}}{900356 - 8 \times (\frac{2672}{8})^2} = \frac{2206}{7908} \approx 0.278958$$

e que $\hat{\sigma}^2 \stackrel{\text{enunc.}}{\approx} 1.415602$, o valor observado da estatística de teste é igual a

$$t = \frac{\hat{\beta}_1 - \beta_{1,0}}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}}} \approx \frac{0.278958 - 0}{\sqrt{\frac{1.415602}{7908}}} \approx 20.849783.$$

Como $t \approx 20.849783 \in W = (-\infty, -2.447) \cup (2.447, +\infty)$ devemos rejeitar H_0 ao n.s. de $\alpha_0 = 5\%$ [assim como a qualquer n.s. superior a 5%. I.e., há forte evidência a favor de H_1 , i.e., a distância percorrida influencia significativamente a quantidade de combustível consumido].