

Duração: 120 minutos

Exame de Época Especial

- As perguntas com múltiplas opções têm apenas uma resposta correta. Caso assinale uma resposta errada ou mais de uma resposta, serão descontados 0.5 valores.
- As respostas às perguntas de escolha única têm de ser assinaladas na folha de instruções distribuída.
- Justificar convenientemente as respostas a todas as restantes perguntas.

Pergunta 1

2 valores

A execução de um projeto de construção de um edifício no tempo programado está relacionada com os seguintes acontecimentos executados a tempo, A = “escavação”, B = “fundação” e C = “superestrutura”, supostamente independentes e com probabilidades iguais a 0.8, 0.7 e 0.9, respetivamente. Calcule a probabilidade de o prazo de execução ser cumprido para a escavação e não ser cumprido em pelo menos uma das outras atividades.

A: 0.567 B: 0.296 C: 0.103 D: 0.056

Acontecimentos e probabilidades:

A = escavação do projeto executada a tempo, $P(A) = 0.8$

B = fundação do projeto executada a tempo, $P(B) = 0.7$

C = superestrutura do projeto executada a tempo, $P(C) = 0.9$

$$\begin{aligned}P[A \cap (\bar{B} \cup \bar{C})] &= P[(A \cap \bar{B}) \cup (A \cap \bar{C})] = P(A \cap \bar{B}) + P(A \cap \bar{C}) - P(A \cap \bar{B} \cap \bar{C}) \\&= P(A) \times [1 - P(B)] + P(A) \times [1 - P(C)] - P(A) \times [1 - P(B)] \times [1 - P(C)] \\&= 0.8 \times 0.3 + 0.8 \times 0.1 - 0.8 \times 0.3 \times 0.1 = 0.296\end{aligned}$$

Pergunta 2

3 valores

Suponha que numa região enchentes ocorrem ao longo do tempo segundo uma lei poissoniana, ou seja, o número de dias com enchentes num ano tem distribuição de Poisson com parâmetro λ . Sabe-se que a probabilidade de não ocorrer nenhum dia com enchente num ano é igual 0.000045.

- (a) Determine o valor de λ . Se não conseguir determiná-lo, use $\lambda = 10$ nas próximas alíneas. (1.0)

X = número de dias com enchentes num ano, $X \sim \text{Poisson}(\lambda)$

$$P(X = 0) = \exp(-\lambda) = 0.000045 \Rightarrow \lambda = -\log(0.000045) = 10.00885$$

- (b) Calcule a probabilidade de não ocorrerem nem mais de 12 nem menos de 7 dias com enchentes num ano. (1.0)

$$P(7 \leq X \leq 12) = F_X(12) - F_X(6) = 0.790717 - 0.129584 = 0.661133$$

$$\text{Se } \lambda = 10, \text{ então } P(7 \leq X \leq 12) = 0.791556 - 0.130141 = 0.661415$$

- (c) Durante um período de 20 anos, em quantos anos se esperaria que ocorressem dois ou mais dias com enchentes? (1.0)

Y = Número de anos em que ocorrem sucesso ($X \geq 2$) em 20 anos.

Supondo que as ocorrências ou não de dois ou mais dias com enchentes em cada ano são ensaios de Bernoulli(p) independentes, com $p = P(X \geq 2) = 1 - F_X(1) = 0.999505$, então

$$Y \sim \text{Binomial}(n = 20, p = 0.999505) \Rightarrow E(Y) = np = 20 \times 0.999505 = 19.99009.$$

$$\text{Se } \lambda = 10, \text{ então } Y \sim \text{Binomial}(n = 20, p = 0.999501) \Rightarrow E(Y) = np = 20 \times 0.999501 = 19.99001.$$

Pergunta 3

2 valores

Sejam X e Y duas variáveis aleatórias (v.a.) indicadoras de características de cada peça produzida por uma dada fábrica, cuja função de probabilidade conjunta é definida pela seguinte tabela:

$X \setminus Y$	-1	0	1
-1	0.23	0.02	0.23
0	0.02	0	0.02
1	0.23	0.02	0.23

Recorrendo às definições e cálculos apropriados, confirme ou não as seguintes afirmações:

1. A função de probabilidade condicional de X dado $Y = 0$ toma o valor 0.5 em $x \in \{-1, 1\}$ e 0, caso contrário.
2. X e Y são v.a. identicamente distribuídas e a distribuição comum é bimodal.
3. X e Y são v.a. dependentes.
4. O valor da função de distribuição conjunta no ponto $(x = 0.5, y = 0)$ é 0.29.

A: 1, 2 e 3 são verdadeiras. B: 2, 3 e 4 são verdadeiras.

C: 1, 3 e 4 são verdadeiras. D: 1, 2 e 4 são verdadeiras.

(X, Y) = par aleatório discreto com função de probabilidade conjunta $P(X = x, Y = y)$ de acordo com a tabela acima.

1. Função de probabilidade de X condicionada em $Y = 0$:

$$P(X = x | Y = 0) = \frac{P(X=x, Y=0)}{P(Y=0)} = \begin{cases} 0.02/0.04 = 0.5, & x = -1, 1 \\ 0, & \text{caso contrário.} \end{cases}$$

2. X e Y são v.a. identicamente distribuídas, i.e.,

$$P(X = x) = \sum_y P(X = x, Y = y) = \begin{cases} 0.48, & x = -1, 1 \\ 0.04, & x = 0 \\ 0, & \text{c.c.} \end{cases} \quad P(Y = y) = \sum_x P(X = x, Y = y) = \begin{cases} 0.48, & y = -1, 1 \\ 0.04, & y = 0 \\ 0, & \text{c.c.} \end{cases}$$

com distribuição comum bimodal em -1 e 1 pois $P(X = -1) = P(X = 1) = \max P(X = x)$ e $P(Y = -1) = P(Y = 1) = \max P(Y = y)$.

3. Como existe pelo menos um (x, y) tal que $P(X = x, Y = y) \neq P(X = x) \times P(Y = y)$, e.g., $P(X = 0, Y = 0) = 0 \neq 0.04 \times 0.04 = P(X = 0) \times P(Y = 0)$, então X e Y são v.a. dependentes.

4. Função de distribuição conjunta de (X, Y) em $(x = 0.5, y = 0)$:

$$F_{X,Y}(0.5, 0) = \sum_{x \leq 0.5} \sum_{y \leq 0} P(X = x, Y = y) = 0.23 + 0.02 + 0.02 + 0 = 0.27.$$

Pergunta 4

3 valores

Admita que o número de veículos que passa numa determinada secção de uma auto-estrada durante um minuto (X) é razoavelmente descrito pela função de densidade de probabilidade

$$f_X(x) = \begin{cases} \frac{x}{100^2}, & 0 \leq x \leq 100, \\ \frac{2}{100} - \frac{x}{100^2}, & 100 \leq x \leq 200, \\ 0, & \text{caso contrário.} \end{cases}$$

- (a) Mostre que o valor esperado e a variância de X são iguais a 100 e $\frac{100^2}{6}$, respetivamente. (1.0)

X = número de veículos que passa numa secção de uma auto-estrada durante um minuto.

$$\begin{aligned} E(X) &= \int_0^{100} x \frac{x}{100^2} dx + \int_{100}^{200} x \left(\frac{2}{100} - \frac{x}{100^2} \right) dx = \frac{1}{100^2} \frac{x^3}{3} \Big|_0^{100} + \frac{2}{100} \frac{x^2}{2} \Big|_{100}^{200} - \frac{1}{100^2} \frac{x^3}{3} \Big|_{100}^{200} \\ &= \frac{100}{3} + 4 \times 100 - 100 - \frac{800}{3} + \frac{100}{3} = 300 - \frac{600}{3} = 100 \\ E(X^2) &= \int_0^{100} x^2 \frac{x}{100^2} dx + \int_{100}^{200} x^2 \left(\frac{2}{100} - \frac{x}{100^2} \right) dx = \frac{1}{100^2} \frac{x^4}{4} \Big|_0^{100} + \frac{2}{100} \frac{x^3}{3} \Big|_{100}^{200} - \frac{1}{100^2} \frac{x^4}{4} \Big|_{100}^{200} \\ &= \frac{100^2}{4} + \frac{2}{100 \times 3} (200^3 - 100^3) - \frac{1}{4 \times 100^2} (200^4 - 100^4) = 100^2 \left(\frac{1}{4} + \frac{14}{3} - \frac{15}{4} \right) = 100^2 \frac{7}{6} \\ Var(X) &= E(X^2) - E(X)^2 = 100^2 \frac{7}{6} - 100^2 = 100^2 \frac{1}{6} \approx 1666.6667 \end{aligned}$$

- (b) Admitindo suposições de trabalho apropriadas, calcule aproximadamente a probabilidade de o número de veículos que passam durante uma hora na referida secção exceder 5500. (2.0)

- X_i : número de veículos que passam no minuto i , $i = 1, \dots, n$, $n = 60$.
- $S_n = \sum_{i=1}^n X_i$, onde X_i são variáveis aleatórias independentes com função densidade de probabilidade $f_X(x)$.

$$E(S_n) = E(\sum_{i=1}^n X_i) = \sum_{i=1}^n E(X_i) \stackrel{i.d.}{=} \times 60 \times 100 = 6000$$

$$Var(S_n) = Var(\sum_{i=1}^n X_i) \stackrel{ind}{=} \sum_{i=1}^n Var(X_i) \stackrel{i.d.}{=} 60 \times 100^2 / 6 = 10^5$$

- Para grandes amostras ($n = 60$), pelo T.L.C., $S_n \stackrel{a}{\sim} N(6000, 10^5)$, i.e., $Z = \frac{S_n - 6000}{\sqrt{10^5}} \stackrel{a}{\sim} N(0, 1)$.

$$P(S_n > 5500) \approx P\left(Z > \frac{5500 - 6000}{\sqrt{10^5}}\right) = P(Z > -1.58) = 1 - F_{N(0,1)}(-1.58) = 1 - 0.0571 \approx 0.9429$$

Pergunta 5

2 valores

Seja $(X_1, \dots, X_n)$ uma amostra aleatória de X a representar a ocorrência ($X = 1$) de um evento de interesse ou não ($X = 0$), com $p = P(X = 1) = 1 - P(X = 0)$. Considere ainda X_1 e $\bar{X}$ (média amostral) como dois estimadores de p .

Indique o valor do quociente entre os erros quadráticos médios de $\bar{X}$ e de X_1 para $n = 10$:

A: 10 **B:** 1 **C:** 0.1 **D:** 0.01

Supondo $X_i \sim \text{Bernoulli}(p)$ independentes, $E(X_i) = p$ e $Var(X_i) = p(1 - p)$, $i = 1, \dots, n$.

$$E(X_1) \stackrel{i.d.}{=} p \quad e \quad Var(X_1) \stackrel{i.d.}{=} p(1 - p)$$

$$E[\bar{X}] = \frac{1}{n} \sum_{i=1}^n E(X_i) \stackrel{i.d.}{=} \frac{n \times E(X)}{n} = E(X) = p \quad e \quad Var[\bar{X}] \stackrel{ind}{=} \frac{1}{n^2} \sum_{i=1}^n Var(X_i) \stackrel{i.d.}{=} \frac{n \times Var(X)}{n^2} = \frac{Var(X)}{n} = \frac{p(1-p)}{n}$$

$$EQM[X_1] = Var[X_1] + (E(X_1) - p)^2 = p(1 - p) \quad e \quad EQM[\bar{X}] = Var[\bar{X}] + (E[\bar{X}] - p)^2 = \frac{p(1-p)}{n}$$

$$e(\bar{X}, X_1) = \frac{EQM[\bar{X}]}{EQM[X_1]} = \frac{p(1-p)/n}{p(1-p)} = \frac{1}{n} = \frac{1}{10} = 0.1$$

Pergunta 6

2 valores

Um fornecedor de fios diz que o seu produto tem uma resistência média que não é inferior a 90 Kgf. Um cliente decide, antes de efetuar qualquer transação, testar aquela afirmação, baseado numa amostra aleatória de 25 fios, usando a informação de que a resistência possui uma distribuição Normal. Se o teste consuzir à aceitação do que diz o fornecedor, o cliente decide comprar o produto.

Sabendo que da amostra analisada se obteve $\bar{x} = 88 \text{ Kgf}$ e $\sum_{i=1}^n (x_i - \bar{x})^2 = 3456 \text{ Kgf}^2$, indique o valor-p do teste que permite decidir se o produto foi comprado.

A: 0.2064 **B:** 0.0536 **C:** 0.4999 **D:** 0.0002

- X : resistência do fio, com pressuposto $X \sim \text{Normal}(\mu, \sigma^2)$.
- $(X_1, \dots, X_n)$ amostra aleatória de X com concretização $n = 25$, $\bar{x} = 88$ e $\sum_{i=1}^n (x_i - \bar{x})^2 = 3456$, e distribuição amostral $\bar{X} \sim \text{Normal}(\mu, \sigma^2/n)$.
- Teste de hipóteses:
 1. Hipóteses: $H_0 : \mu \geq 90 \equiv \mu_0$ versus $H_1 : \mu < 90$.
 2. Estatística do teste: $T = \frac{\bar{X} - \mu_0}{\sqrt{S^2/n}} \underset{\mu=\mu_0}{\sim} t_{(24)}$, com valor observado $t = \frac{88-90}{\sqrt{144/25}} \approx -0.8333$, visto que $s^2 = 3456/24 = 144$.
 3. Valor-p: $P = P(T < -0.8333 | H_0) = F_{t_{(24)}}(-0.8333) = 0.2064$. Note-se que $P \in (0.2 = F_{t_{(24)}}(-0.857), 0.25 = F_{t_{(24)}}(-0.685))$.
 4. Conclusão: Rejeita-se H_0 para $\alpha > 0.2023$ e aceita-se H_0 no caso contrário. Portanto, tendo em conta os níveis de significância usuais 1%, 5% e 10%, aceita-se H_0 e o cliente deve decidir pela compra do produto.

Pergunta 7 (corrigida)

3 valores

Com o objetivo de se estudar a variabilidade do tempo de vida X (em horas) das válvulas produzidas por uma dada empresa, ensaiaram-se 1200 válvulas anotando-se os seus tempos de vida. O agrupamento destes tempos com vista a testar a hipótese de X ter distribuição Exponencial com $E(X) = 28$ horas está reproduzido na tabela abaixo:

classe	]0, 20[	]20, 40[	]40, 60[	]60, ∞ [
frequência observada (o_i)	600	300	170	130
frequência esperada (e_i)	-	-	146.80	140.78

- (a) Determine os valores das frequências esperadas em falta na tabela.

(1.0)

$X \sim \text{Exponencial}(\lambda)$ com $E(X) = 1/\lambda = 28 \Rightarrow \lambda = 1/28$.

Função de distribuição de X : $F_X(x) = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x} = 1 - e^{-x/28}$, $x > 0$.

$F_X(20) = 1 - e^{-20/28} = 0.510458 \Rightarrow e_1 = 1200 \times 0.510458 \approx 612.55$.

$F_X(40) = 1 - e^{-40/28} = 0.760349 \Rightarrow e_2 = 1200 \times (0.760349 - 0.510458) = 1200 \times 0.249891 \approx 299.87$.

- (b) Que conclusões se pode retirar usando estes dados relativamente à hipótese de interesse?

(2.0)

- Teste de ajustamento do Qui-quadrado de Pearson com $k = 4$ classes.
 1. Hipóteses: $H_0 : X \sim \text{Exponencial}(\lambda = 1/28)$ versus $H_1 : X \not\sim \text{Exponencial}(\lambda = 1/28)$.
 2. Estatística do teste: $T = \sum_{i=1}^4 \frac{(O_i - E_i)^2}{E_i} \underset{H_0}{\sim} \chi_{(k-1)}^2 = \chi_{(3)}^2$.

i	classe $(x)_i$	o_i	$e_i = 1200 \times p_i^0$	$\frac{(o_i - e_i)^2}{e_i}$
1	]0, 20[	600	612.55	0.257126
2	[20, 40[	300	299.87	0.000057
3	[40, 60[	170	146.80	3.667084
4	[60, $+\infty$ [	130	140.78	0.825903
total		1200	1200	$t \approx 4.75017$

3. Valor-p: $P = P(T \geq 4.75017 | H_0) \approx 1 - F_{\chi_{(3)}^2}(4.75017) = 1 - 0.809 = 0.191$. Note-se que $P \in (0.15, 0.2) = (1 - F_{\chi_{(3)}^2}(5.317), 1 - F_{\chi_{(3)}^2}(4.642))$.

4. Conclusão: Como o valor-p é elevado (> 0.10), não se rejeita a hipótese nula H_0 para os níveis de significância usuais ($\alpha = 1\%, 5\%$ ou 10%), isto é, com base nesta amostra não existem evidências para rejeitar a adequação do modelo.

Pergunta 8

3 valores

Visando investigar a resistência de molas, registou-se respetivamente a carga aplicada (x) e a deformação sofrida (y) em 8 molas, tendo-se obtido os seguintes resultados:

x	11.2	21.1	29.9	34.1	43.8	59.9	61.2	68.9	$\sum_{i=1}^8 x_i = 330.1, \sum_{i=1}^8 y_i = 31.4, \sum_{i=1}^8 x_i y_i = 1531.46$
y	1.6	2.1	3.4	3.3	4.2	4.3	6.2	6.3	$\sum_{i=1}^8 x_i^2 = 16626.57, \sum_{i=1}^8 y_i^2 = 143.68$

Considere um modelo de regressão linear simples para explicar as observações.

- (a) Obtenha a estimativa de mínimos quadrados dos coeficientes de reta de regressão. (1.5)

Modelo e hipóteses de trabalho: $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, i = 1, \dots, 12$, com $\varepsilon_i \stackrel{iid}{\sim} N(0, \sigma^2)$.

Estimadores de mínimos quadrados de β_0 e β_1 :

$$\hat{\beta}_1 = \frac{\sum_{i=1}^8 y_i x_i - n \bar{y} \bar{x}}{\sum_{i=1}^8 x_i^2 - n \bar{x}^2} = \frac{1531.46 - 8 \times 3.925 \times 41.2625}{16626.57 - 8 \times 41.2625^2} = 0.078454$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} = 3.925 - 0.078454 \times 41.2625 = 0.687806$$

- (b) Calcule o intervalo de confiança a 95% para o verdadeiro valor do declive da reta de regressão. (1.5)

- Método da variável fulcral (estimação intervalar de β_1):

1. Variável fulcral:

$$T = \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}}} \sim t_{(n-2)} = t_{(6)}$$

2. Fixado o grau de confiança $1 - \alpha = 0.95$, obtêm-se a e b tais que

$$P(a < T < b) = 1 - \alpha = 0.95 \Rightarrow a = -b \text{ e } b = F_{t_{(6)}}^{-1}(0.975) = 2.4469$$

3. Desigualdade equivalente:

$$P\left(-b < \frac{\hat{\beta}_1 - \beta_1}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}}} < b\right) = 0.95$$

$$\Leftrightarrow P\left(\hat{\beta}_1 - b\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}} < \beta_1 < \hat{\beta}_1 + b\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}}\right) = 0.95,$$

4. Intervalo aleatório de confiança (IAC) e sua concretização (IC) a 95% para β_1 :

$$IAC(\beta_1; 0.95): \hat{\beta}_1 \pm b\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}}$$

$$IC(\beta_1; 0.95): (0.078454 \pm 2.4469 \times 0.010356) \\ (0.053113, 0.103794)$$

tendo em conta os cálculos auxiliares abaixo.

$$\hat{\sigma}^2 = \frac{1}{n-2} \left[\left(\sum_{i=1}^8 y_i^2 - n\bar{y}^2 \right) - (\hat{\beta}_1)^2 \left(\sum_{i=1}^8 x_i^2 - n\bar{x}^2 \right) \right] = \frac{1}{6} [20.435 - (0.078454)^2 3005.819] = 0.322375$$

$$\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}} = \sqrt{\frac{0.322375}{3005.819}} \approx 0.010356.$$