

Duração: **120** minutos

Exame de Época de Recurso - Versão F

- 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

Numa determinada especialidade de um hospital, os doentes são classificados em uma de três classes C_1 , C_2 ou C_3 , consoante o maior ou menor risco que se lhes atribui. A probabilidade de um doente ao acaso ser classificado em C_1 , C_2 ou C_3 é igual a 0.2, 0.4, 0.4, respetivamente. A probabilidade do doente ser atendido rapidamente, estando classificado como sendo de C_1 , C_2 ou C_3 é de 0.6, 0.7 e 0.9, respetivamente. Qual a classificação mais provável de um doente que não foi atendido rapidamente?

A: C_1 **B:** C_2 **C:** C_3 **D:** C_2 ou C_3

Acontecimentos e probabilidades:

R = doente atendido rapidamente

C_1 = doente classificado na classe C_1 , $P(C_1) = 0.2$, $P(R|C_1) = 0.6$

C_2 = doente classificado na classe C_2 , $P(C_2) = 0.4$, $P(R|C_2) = 0.7$

C_3 = doente classificado na classe C_3 , $P(C_3) = 0.4$, $P(R|C_3) = 0.9$

Pela Lei da Probabilidade Total,

$$P(R) = P(R|C_1)P(C_1) + P(R|C_2)P(C_2) + P(R|C_3)P(C_3) = 0.76$$

Logo,

$$\begin{aligned} P(C_1|\bar{R}) &= \frac{(1-P(R|C_1))P(C_1)}{P(\bar{R})} = \frac{0.4 \cdot 0.2}{0.24} = 1/3 \\ P(C_2|\bar{R}) &= \frac{(1-P(R|C_2))P(C_2)}{P(\bar{R})} = 0.5 \\ P(C_3|\bar{R}) &= \frac{(1-P(R|C_3))P(C_3)}{P(\bar{R})} = 1/6 \end{aligned}$$

Pergunta 2

2 valores

Considere que o número de contentores movimentados por hora num terminal portuário de dimensão média é uma variável aleatória X com distribuição de Poisson. Sabendo que $P(X = 100) = 1.006 \times P(X = 99)$, determine a diferença entre o valor esperado e a moda de X .

A: 3.5 **B:** 0.4 **C:** 1.5 **D:** 0.6

$$P(X = 100) = 1.006 \times P(X = 99) \iff e^{-\lambda} \frac{\lambda^{100}}{100!} = 1.006 \times e^{-\lambda} \frac{\lambda^{99}}{99!} \iff E[X] = \lambda = 100.6$$

Seja $g(x) = \frac{f_X(x+1)}{f_X(x)} = \frac{100.6}{x+1}$, $\forall x \in \mathbb{N}^0$. Como $g(x) > 1$ para $x \leq 99$ e $g(x) < 1$ para $x \geq 100$, conclui-se que Moda (X) = 100 e, assim, $E[X] - \text{Moda}(X) = 0.6$.

Pergunta 3

3 valores

Admita que, durante a sua vida útil, um determinado componente eletrónico apresenta uma taxa de falha constante ao longo do tempo. Nessas condições, a duração, em horas, do componente é modelado por uma variável aleatória contínua X com distribuição exponencial de parâmetro $\lambda = 0.05$.

- (a) Sabendo que o componente já funcionou durante 10 horas, determine a probabilidade de ele durar (1.5)

pelo menos mais 15 horas.

Como X tem distribuição exponencial ($\lambda = 0.05$), verifica-se a propriedade de perda de memória:

$$P(X > s + t \mid X > s) = P(X > t).$$

Logo,

$$P(X > 25 \mid X > 10) = P(X > 15) = 1 - P(X \leq 15) = 1 - \int_0^{15} \lambda e^{-\lambda t} dt = e^{-0.05 \times 15} = e^{-0.75} \approx 0.4724.$$

(b) Determine a duração mediana da componente eletrônica.

(1.5)

Função de distribuição de X : $F_X(x) = \int_0^x \lambda e^{-\lambda t} dt = e^{-\lambda t} \Big|_0^x = 1 - e^{-\lambda x}$, para $x > 0$.

Mediana de X (x_d): $F_X(x_d) = 0.5$, i.e., $x_d = -\log(1 - 0.5)/\lambda = 13.86294$.

Pergunta 4

3 valores

Num reator de fluxo contínuo, a temperatura (T) e o rendimento (R) são modelados como variáveis aleatórias contínuas. Devido à natureza exotérmica da reação, estas variáveis não são independentes. A função de densidade de probabilidade conjunta, para valores normalizados de T e R no intervalo $[0, 1]$, é dada por:

$$f_{T,R}(t, r) = \begin{cases} \frac{3}{4}(t^2 + 2r), & 0 \leq t \leq 1 \text{ e } 0 \leq r \leq 1 \\ 0, & \text{caso contrário} \end{cases}$$

Tenha em conta que $E[T] = \frac{3}{2} \left(E[R] - \frac{2}{8} \right)$ e que $E[TR] = \frac{1}{4} \left(E[R] + \frac{6}{8} \right)$.

(a) Calcule o valor da covariância entre T e R .

(2.0)

Tendo em conta as relações dadas temos que $Cov[T, R] = E[TR] - E[T]E[R] = -\frac{3}{2}(E[R])^2 + \left(\frac{1}{4} + \frac{3}{2}\frac{2}{8}\right)E[R] + \frac{1}{4} \times \frac{6}{8}$

$$E[R] = \int_0^1 \int_0^1 r \frac{3}{4}(t^2 + 2r) dr dt = \frac{3}{4} \left(\frac{1}{6} + \frac{2}{3} \right) = 0.625$$

$$Cov[T, R] = -0.0078125$$

(b) Serão T e R variáveis aleatórias independentes?

(1.0)

Se T e R são variáveis aleatórias independentes, então $Cov[T, R] = 0$. Como $Cov[T, R] \neq 0$, as variáveis aleatórias T e R não são independentes.

Pergunta 5

3 valores

Um engenheiro de propulsão está a analisar a duração X de válvulas de combustível de um foguetão (em horas) e admite que essa variável aleatória segue uma distribuição exponencial de parâmetro $\lambda \in \mathbb{R}^+$.

- (a) Determine o estimador de máxima verosimilhança de λ com base numa amostra aleatória de dimensão n de X . (2.0)

Seja $\mathbf{x} = (x_1, \dots, x_n)$.

$$\mathcal{L}(\lambda | \mathbf{x}) = \lambda^n \exp(-\lambda \sum_{i=1}^n x_i)$$

$$\log \mathcal{L}(\lambda | \mathbf{x}) = n \log \lambda - \lambda \sum_{i=1}^n x_i$$

$$\frac{d \log \mathcal{L}(\lambda | \mathbf{x})}{d\lambda} = 0 \iff \frac{n}{\lambda} - \sum_{i=1}^n x_i = 0 \iff \lambda = \frac{n}{\sum_{i=1}^n x_i}$$

$$\frac{d^2 \log \mathcal{L}(\lambda | \mathbf{x})}{d\lambda^2} = -\frac{n}{\lambda^2} < 0, \forall \lambda \in \mathbb{R}^+$$

$$\therefore \hat{\lambda}_{MV} = \frac{n}{\sum_{i=1}^n x_i}$$

- (b) Para uma amostra de 25 válvulas em que a soma das durações observadas foi de 9972 horas, calcule a estimativa da probabilidade de uma válvula de combustível durar mais do que 500 horas. (1.0)

Seja $g(\lambda) = P(X > 500) = 1 - F_X(500) = e^{-500\lambda}$. Pela invariância dos EMV, tem-se $\hat{g}_{MV}(\lambda) = e^{-500\hat{\lambda}_{MV}} = e^{-500 \times \frac{25}{9972}} = 0.285501$

Pergunta 6

3 valores

Numa exploração mineira, a viabilidade de processar um determinado setor depende da proporção de blocos de minério que apresentam um teor de metais acima de 60%. Para estimar essa proporção, o engenheiro de minas analisou uma amostra de 200 blocos extraídos de forma aleatória. Após a análise laboratorial, verificou-se que 122 blocos tinham um teor de metais superior a 60%.

- (a) Determine um intervalo de confiança a aproximadamente 98% para a proporção populacional de blocos com teor de metais superior a 60% naquele setor da mina. (1.5)

- Variável aleatória de interesse e distribuição:
 X , tal que $X = 1$ se teor de metais $> 60\%$ e $X = 0$ se teor de metais $\leq 60\%$, $X \sim \text{Bernoulli}(p)$
- Método da variável fulcral (estimação intervalar de p):
 1. Variável fulcral (aproximada):

$$Z = \frac{\hat{P} - p}{\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}} \stackrel{a}{\sim} N(0, 1)$$

aproximação válida para n grande, o que se verifica ($n = 200$).

2. Fixado o grau de confiança $1 - \alpha = 0.98$, obtêm-se a e b tais que

$$P(a < Z < b) = 1 - \alpha = 0.98 \text{ e } a = -b \Rightarrow b = \Phi^{-1}(0.99) = 2.3263$$

3. Desigualdade equivalente:

$$P\left(-b < \frac{\hat{P} - p}{\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}} < b\right) = 0.98 \Leftrightarrow P\left(\hat{P} - b\sqrt{\frac{\hat{P}(1-\hat{P})}{n}} < p < \hat{P} + b\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}\right) = 0.98$$

4. Intervalo aleatório de confiança (IAC) e sua concretização (IC) a 98% para p :

$$IAC(p; 0.98) : \hat{P} \pm b\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$$

$$IC(p; 0.98) : 0.61 \pm 2.3263 \times \sqrt{\frac{0.61 \times (1 - 0.61)}{200}} \approx (0.53; 0.69),$$

tendo em conta que $\hat{p} = 122/200 = 0.61$

- (b) Suponha que a administração da mina exige uma estimativa mais precisa para decidir sobre o investimento em nova maquinaria. Qual deve ser o tamanho mínimo da amostra para que a margem de erro da proporção não exceda 0.03, com o mesmo nível de confiança aproximado de 98%? (Considere a proporção observada, 0.61, como uma estimativa preliminar). (1.5)

- A margem de erro da proporção é $b\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$ (da alínea anterior). Como o nível de confiança é o mesmo (98%), $b = 2.3263$
- Pretende-se determinar n tal que

$$b\sqrt{\frac{\hat{P}(1-\hat{P})}{n}} \leq 0.03 \Leftrightarrow n \geq \frac{b^2 \hat{P}(1-\hat{P})}{0.03^2}$$

- Usando a estimativa preliminar $\hat{p} = 122/200 = 0.61$,

$$n \geq \frac{2.3263^2 \times 0.61 \times (1 - 0.61)}{0.03^2} \approx 1430.544$$

- O tamanho mínimo da amostra que verifica as condições impostas é 1431.

Pergunta 7

2 valores

A resistência à tração de amostras de certo tipo de madeira, medido em unidades apropriadas, X , possui distribuição normal de valor esperado μ , desconhecido, e variância $\sigma^2 = 4$.

Determine o valor-p do teste $H_0 : \mu = 20$ versus $H_1 : \mu \neq 20$, sabendo que uma amostra de tamanho 16 conduziu a $\bar{x} = 19$.

A: 0.0266 B: 0.0456 C: 0.0522 D: 0.0744

- Estatística de teste:

$$T = \frac{\bar{X} - 20}{\sigma/\sqrt{16}} = 2(\bar{X} - 20) \underset{H_0}{\sim} \text{Normal}(0, 1).$$

- Valor observado da estatística de teste: $t = 2(\bar{x} - 20) = 2(19 - 20) = -2$.
- Valor-p = $P(|T| > |t| | H_0) = 2[1 - \Phi(2)] = 2[1 - 0.9772] = 2 \times 0.1587 = 0.0456$.

Pergunta 8

2 valores

No departamento de controlo de qualidade de uma empresa metalúrgica, avalia-se o diâmetro de um novo tipo de parafuso (X , em mm). Com base no histórico de produção, conjectura-se que o diâmetro segue uma distribuição Normal com média $\mu = 10.0$ e desvio-padrão $\sigma = 0.1$. Para testar esta conjectura, recolheu-se uma amostra de $n = 300$ parafusos, cujas medições produziram os seguintes resultados:

Diâmetro (X)	$] -\infty, 9.9[$	$[9.9, 10.1[$	$[10.1, +\infty[$
Frequência observada (o_i)	52	192	56

Averigue se o modelo $N(10, 0.1^2)$ é adequado para descrever a variação do diâmetro dos parafusos. Decida com base no valor-p aproximado.

A: Valor-p < 0.01, não se rejeita H_0 B: Valor-p < 0.01, rejeita-se H_0
 C: Valor-p > 0.10, não se rejeita H_0 D: Valor-p > 0.10, rejeita-se H_0

- $X \sim N(10, 0.1^2)$, a sua função de distribuição é $F_X(x) = \Phi\left(\frac{x-10}{0.1}\right)$.
- Teste de ajustamento do Qui-quadrado de Pearson com $k = 3$ classes.

1. Hipóteses: $H_0 : X \sim N(10, 0.1^2)$ versus $H_1 : X \not\sim N(10, 0.1^2)$.

As probabilidades de cada classe, sob H_0 , são $p_i^0 = P(X \in \text{classe}_i) = F_X(\text{lim.sup}) - F_X(\text{lim.inf})$:

- $p_1^0 = P(X < 9.9) = \Phi\left(\frac{9.9-10}{0.1}\right) = \Phi(-1) = 0.1587$
- $p_2^0 = P(9.9 \leq X < 10.1) = \Phi(1) - \Phi(-1) = 0.8413 - 0.1587 = 0.6826$
- $p_3^0 = P(X \geq 10.1) = 1 - \Phi(1) = 1 - 0.8413 = 0.1587$

2. Estatística do teste: $T = \sum_{i=1}^3 \frac{(O_i - E_i)^2}{E_i} \underset{H_0}{\sim} \chi_{(k-1)}^2 = \chi_{(2)}^2$.

1	$] -\infty, 9.9[$	52	47.597	0.4074
2	$[9.9, 10.1[$	192	204.807	0.8008
3	$[10.1, +\infty[$	56	47.596	1.4837
total		300	300	$t \approx 2.6919$

3. Valor-p: $P = P(T \geq 2.6919 | H_0) \approx 1 - F_{\chi_{(2)}^2}(2.6919) = 1 - 0.7398 = 0.2602$. Note-se que $P \in (0.2, 0.3) = (1 - F_{\chi_{(2)}^2}(3.219), 1 - F_{\chi_{(2)}^2}(2.408))$.

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.