

Duração: 120 minutos

Exame de Época de Recurso - Versão E

- 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

3 valores

Num processo de recrutamento tem-se verificado que 40% dos candidatos passam na primeira fase. Dos que passam na primeira fase, 60% passam na segunda fase. Finalmente dos que chegam à terceira fase, 70% são selecionados.

- (a) Qual é a probabilidade de um candidato ser selecionado?

(1.5)

Acontecimentos e probabilidades:

S = candidato selecionado = $F_1 \cap F_2 \cap F_3$

F_1 = candidato passar na primeira fase, $P(F_1) = 0.4$

F_2 = candidato passar na segunda fase, $P(F_2|F_1) = 0.6$

F_3 = candidato passar na terceira fase, $P(F_3|F_1 \cap F_2) = 0.7$

$$P(S) = P(F_1 \cap F_2 \cap F_3) = P(F_3|F_1 \cap F_2) \times P(F_2|F_1) \times P(F_1) = 0.168$$

- (b) Suponha que todos os candidatos selecionados são contratados e que, adicionalmente, 20% dos candidatos que não passam na primeira fase são repescados e acabam por ser contratados. Qual é a probabilidade de um candidato ser contratado?

(1.5)

Acontecimentos e probabilidades:

C = candidato contratado

$P(C|S) = 1$, $P(C|\bar{F}_1) = 0.2$

Pela Lei da Probabilidade Total,

$$P(C) = P(C|S)P(S) + P(C|\bar{F}_1)P(\bar{F}_1) = 1 \times P(S) + P(C|\bar{F}_1)P(\bar{F}_1) = 0.168 + 0.2 \times 0.6 = 0.288$$

Pergunta 2

2 valores

Na monitorização da estabilidade de taludes em minas a céu aberto, utilizam-se inclinómetros para detetar movimentos no terreno. Se um sensor tem uma precisão nominal de 1° , o erro de leitura L pode ser modelado como $L \sim \text{Uniforme}(-1, 1)$. Nestas condições, indique o valor da variância de $3L^2 + 1$:

A: 0.41 **B:** 0.80 **C:** 0.72 **D:** 0.32

$$\begin{aligned} \text{Var}[3L^2 + 1] &= 9\text{Var}[L^2] = 9(E[L^4] - (E[L^2])^2) = 9(E[L^4] - (E[L^2])^2) = \\ &= 9(E[L^4] - (\text{Var}[L] + (E[L])^2)^2) = 9(E[L^4] - (\frac{1}{3} + 0)^2) \end{aligned}$$

$$E[L^4] = \int_{-1}^1 l^4 \frac{1}{2} dl = \frac{1}{5}$$

$$\text{Var}[3L^2 + 1] = 9(\frac{1}{5} - \frac{1}{9}) = 0.8$$

Pergunta 3

2 valores

Um sensor industrial é submetido a choques sucessivos e independentes. A probabilidade do sensor avariar em cada choque é 0.2. Seja X o número de choques até ocorrer a primeira avaria. Sabendo que o sensor já resistiu a 4 choques, determine a probabilidade de ele resistir a pelo menos mais 2 choques.

A: 0.04 B: 0.98 C: 0.36 D: 0.64

V.a.: X = número de choques até ocorrer a primeira falha.

Distribuição: Como a probabilidade de falha em cada choque é constante e igual a $p = 0.2$, e os choques são independentes, tem-se $X \sim \text{geométrica}(0.2)$.

Pretende-se: calcular a probabilidade de o sensor resistir a pelo menos mais 2 choques, sabendo que já resistiu a 4 choques, isto é,

$$P(X > 6 \mid X > 4).$$

Perda de memória: Como X tem distribuição geométrica, verifica-se a propriedade de perda de memória:

$$P(X > m + n \mid X > m) = P(X > n). \quad \text{Assim,} \quad P(X > 6 \mid X > 4) = P(X > 2).$$

Para a distribuição geométrica, tem-se

$$P(X > n) = (1 - p)^n. \quad \text{Logo,} \quad P(X > 2) = (1 - 0.2)^2 = 0.8^2 = 0.64.$$

Conclusão: A probabilidade de o sensor resistir a pelo menos mais 2 choques é 0.64.

Pergunta 4

3 valores

A capacidade de um anfiteatro numa determinada faculdade é de 150 alunos. Tendo em conta a experiência passada, sabe-se que apenas 35% dos alunos inscritos em qualquer disciplina frequentam as aulas.

Para uma aula de uma disciplina com 350 alunos inscritos calcule, recorrendo ao teorema do limite central, a probabilidade do número de alunos presentes exceder a capacidade do anfiteatro.

- X_i : variável indicadora se o caloiro i frequenta as aulas, $i = 1, \dots, n$.
- $S_n = \sum_{i=1}^n X_i$, onde X_i são variáveis aleatórias Bernoulli ($p = 0.35$) independentes e $n = 350$.

$$E(S_n) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) \stackrel{i.d.}{=} n \times 0.35 = 122.5$$

$$Var(S_n) = Var\left(\sum_{i=1}^n X_i\right) \stackrel{ind}{=} \sum_{i=1}^n Var(X_i) \stackrel{i.d.}{=} n \times 0.35(1 - 0.35) = 79.625$$

- Para grandes amostras ($n = 400$), pelo T.L.C., $S_n \stackrel{d}{\sim} N(122.5, 79.625)$, i.e., $Z = \frac{S_n - 122.5}{\sqrt{79.625}} \stackrel{d}{\sim} N(0, 1)$.

$$P(S_n > 150) = P\left(Z > \frac{150 - 122.5}{\sqrt{79.625}}\right) = P(Z > 3.0818) \approx 1 - F_{N(0,1)}(3.0818) = 1 - 0.999 = 0.001$$

Pergunta 5

3 valores

Um engenheiro biomédico está a calibrar um detetor de partículas numa máquina de *PET Scan*. Para tal, mede o número de partículas detetadas em intervalos de um segundo e admite que o número de partículas detetadas por segundo, X , segue uma distribuição de Poisson com 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 $\mathbf{x} = (x_1, \dots, x_n)$.

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

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

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

$$\frac{d^2 \log \mathcal{L}(\lambda | \mathbf{x})}{d\lambda^2} = -\frac{\sum_{i=1}^n x_i}{\lambda^2} < 0, \forall \lambda \in \mathbb{R}^+ \text{ e } \forall \mathbf{x} : \sum_{i=1}^n x_i > 0$$

$$\therefore \hat{\lambda}_{MV} = \bar{X}$$

- (b) Tendo em conta que em 35 intervalos de um segundo, o engenheiro registou um total de 182 partículas detetadas, calcule a correspondente estimativa da probabilidade de serem detetadas mais do que 2 partículas num segundo. (1.0)

Tem-se $\hat{\lambda}_{MV} = \bar{x} = 5.2$. Seja $g(\lambda) = P(X > 2) = 1 - F_X(2)$. Pela invariância dos EMV, tem-se $\hat{g}_{MV}(\lambda) = 1 - F_{Poi(\hat{\lambda}_{MV})}(2) = 1 - F_{Poi(5.2)}(2) = 0.89121$

Pergunta 6

2 valores

Uma equipa de engenharia do ambiente está a avaliar a concentração de mercúrio nos efluentes de uma unidade industrial antes da sua descarga num rio local. Sabe-se, pelos registos históricos, que a concentração de mercúrio, em $\mu g/L$, naquele tipo de efluentes segue uma distribuição normal. Para controlar a conformidade com as normas ambientais, foram efetuadas 12 medições independentes num período de 24 horas, tendo-se obtido os seguintes valores:

$$\sum_{i=1}^{12} x_i = 149.4 \quad \text{e} \quad \sum_{i=1}^{12} x_i^2 = 1863.42$$

Calcule um intervalo de confiança a 90% para o desvio padrão da concentração de mercúrio nos efluentes desta fábrica.

A: (0.49, 0.87) **B:** (0.39, 0.87) **C:** (0.42, 0.86) **D:** (0.39, 0.97)

- Variável aleatória de interesse e distribuição:

X = concentração de mercúrio no efluente da fábrica, $X \sim N(\mu, \sigma^2)$

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

1. Variável fulcral:

$$Q = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)} = \chi^2_{(11)}$$

2. Fixado o grau de confiança $1 - \alpha = 0.90$, obtêm-se a e b tais que

$$P(a < Q < b) = 1 - \alpha = 0.9 \text{ e } P(Q < a) = P(Q > b) \Rightarrow a = F_{\chi^2_{(11)}}^{-1}(0.05) = 4.575 \text{ e } b = F_{\chi^2_{(11)}}^{-1}(0.95) = 19.675$$

3. Desigualdade equivalente:

$$P\left(a < \frac{(n-1)S^2}{\sigma^2} < b\right) = 0.9 \Leftrightarrow P\left(\frac{S\sqrt{n-1}}{\sqrt{b}} < \sigma < \frac{S\sqrt{n-1}}{\sqrt{a}}\right) = 0.9$$

4. Intervalo aleatório de confiança (IAC) e sua concretização (IC) a 90% para σ :

$$IAC(\sigma; 0.9) : \left(\frac{S\sqrt{n-1}}{\sqrt{b}}, \frac{S\sqrt{n-1}}{\sqrt{a}} \right)$$

$$IC(\sigma; 0.9) : \left(\frac{0.555141\sqrt{11}}{\sqrt{19.675}}, \frac{0.555141\sqrt{11}}{\sqrt{4.575}} \right) \approx (0.42, 0.86),$$

tendo em conta os cálculos auxiliares abaixo.

$$\bar{x} = \frac{\sum_{i=1}^{12} x_i}{n} = \frac{149.4}{12} = 12.45$$

$$s = \sqrt{\frac{\sum_{i=1}^{12} x_i^2 - n\bar{x}^2}{n-1}} = \sqrt{\frac{1863.42 - 12 \times 12.45^2}{11}} = 0.555141$$

Pergunta 7

2 valores

A quantidade de energia, X em kW, consumida diariamente numa empresa possui distribuição normal de média μ e variância σ^2 , ambos desconhecidos.

Sabendo que 21 observações recolhidas ao acaso conduziram a $\bar{x} = 105$ e $s^2 = 8$, o teste das hipóteses $H_0 : \sigma^2 = 10$ versus $H_1 : \sigma^2 < 10$ permite concluir que:

A: Valor-p = 0.72 e, portanto, não se rejeita H_0 **B:** Valor-p = 0.72 e, portanto, rejeita-se H_0

C: Valor-p = 0.28 e, portanto, rejeita-se H_0 **D:** Valor-p = 0.28 e, portanto, não se rejeita H_0

- Estatística de teste:

$$Q = \frac{(21-1)S^2}{10} \sim_{H_0} \chi^2_{(20)}.$$

- Valor observado da estatística de teste: $q = \frac{(21-1) \times 8}{10} = 16$

- Valor-p: $P = P(Q < q | H_0) = F_{\chi^2_{(20)}}(16) = 0.2834$

- Conclusão: Rejeita-se H_0 para níveis de significância $\alpha > P = 0.2834$ e não se rejeita H_0 no caso contrário. Ou seja, não se rejeita H_0 ao nível de significância usuais (1%, 5% e 10%).

Um engenheiro eletrotécnico está a estudar a relação entre a temperatura ambiente (x , em $^{\circ}\text{C}$) e a resistência elétrica (Y , em $m\Omega$) num segmento de um condutor de calor. Admite-se que a resistência pode variar linearmente com a temperatura.

Foi recolhida uma amostra com 12 medições independentes, tendo-se obtido os seguintes resultados sumários:

$$\sum_{i=1}^{12} x_i = 300 \quad \sum_{i=1}^{12} y_i = 144 \quad \sum_{i=1}^{12} x_i^2 = 8000 \quad \sum_{i=1}^{12} y_i^2 = 1750 \quad \sum_{i=1}^{12} x_i y_i = 3680$$

Considere o modelo de regressão linear simples, $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, e efetue um teste de hipóteses para averiguar se a temperatura tem um efeito estatisticamente significativo na resistência do condutor.

Utilize um nível de significância $\alpha = 0.01$ e admita as hipóteses de trabalho convenientes.

- Modelo e hipóteses de trabalho: $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $i = 1, \dots, 12$, com $\varepsilon_i \stackrel{iid}{\sim} N(0, \sigma^2)$
- Hipóteses em teste: $H_0 : \beta_1 = 0$ vs. $H_1 : \beta_1 \neq 0$ (teste bilateral)
- Estatística de teste:

$$T_0 = \frac{\hat{\beta}_1}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}}} \stackrel{H_0}{\sim} t_{(n-2)} = t_{(10)}$$

- Região crítica: (Rejeição de H_0) $\{|T_0| > a\}$ com $a = F_{t_{(10)}}^{-1}(1 - 0.01/2) = F_{t_{(10)}}^{-1}(0.995) = 3.169$
- Valor observado da estatística de teste:

$$t_0 = \frac{\hat{\beta}_1}{\sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}}} = \frac{0.16}{\sqrt{\frac{0.92}{8000 - 12 \times 25^2}}} \approx \frac{0.16}{0.0429} \approx 3.73$$

cálculos auxiliares:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{300}{12} = 25 \quad \bar{y} = \frac{\sum_{i=1}^n y_i}{n} = \frac{144}{12} = 12 \quad \hat{\beta}_1 = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^n x_i^2 - n\bar{x}^2} = \frac{3680 - 12 \times 25 \times 12}{8000 - 12 \times 25^2} = 0.16$$

$$\hat{\sigma}^2 = \frac{1}{n-2} \left[\left(\sum_{i=1}^n y_i^2 - n\bar{y}^2 \right) - (\hat{\beta}_1)^2 \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right) \right] = \frac{(1750 - 12 \times 12^2) - 0.16^2 \times (8000 - 12 \times 25^2)}{10} = 0.92$$

- Decisão: como $|t_0| > a$, rejeitamos H_0 ao nível de significância de 1%.
- Conclusão: há forte evidência estatística, ao nível de significância de 1%, de que a temperatura ambiente tem efeitos na resistência do condutor.