

Duração: 120 minutos

Exame de Época Normal - Versão C

- 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

Num dado equipamento constituído por 10 transístores do mesmo tipo, em que 8 deles estão em perfeitas condições, são retirados ao acaso, sucessivamente e sem reposição, 2 transístores.

Determine a probabilidade de o primeiro transistor retirado estar em perfeitas condições sabendo que o segundo transistor retirado não está em perfeitas condições.

A: 9/10 B: 9/11 C: 16/17 D: 8/9

Acontecimentos e probabilidades:

F_i = i -ésimo transistor retirado está em perfeitas condições, $i = 1, 2$.

$$P(F_1|\bar{F}_2) = \frac{P(F_1) \times P(\bar{F}_2|F_1)}{P(F_1) \times P(\bar{F}_2|F_1) + P(\bar{F}_1) \times P(\bar{F}_2|\bar{F}_1)} = \frac{\frac{8}{10} \times \frac{2}{9}}{\frac{8}{10} \times \frac{2}{9} + \frac{2}{10} \times \frac{1}{9}} = \frac{8}{9}$$

Pergunta 2

3 valores

Considere que cada sensor de qualidade numa linha de produção identifica corretamente uma peça defeituosa com probabilidade p operando de modo independente dos demais sensores. Seja X a variável aleatória que representa o número de sensores que identificam corretamente uma peça defeituosa em n sensores utilizados na inspeção da peça defeituosa. Suponha que:

- $E(X) = 6$
- $Var(X) = 1.5$

(a) Qual a distribuição de probabilidade de X ? Especifique os seus parâmetros.

(1.5)

X : o número de sensores que identificam corretamente uma peça defeituosa em n sensores.

Como cada sensor identifica corretamente a peça defeituosa com probabilidade p , independentemente dos restantes sensores, têm-se n ensaios de Bernoulli(p) independentes $\Rightarrow X \sim \text{Binomial}(n, p)$.

Sabe-se que $E(X) = np = 6$ e $Var(X) = np(1 - p) = 1.5$, pelo que

$$Var(X) = 1.5 = 6(1 - p) \Rightarrow p = 1 - \frac{1.5}{6} = 0.75 \quad \text{e} \quad E(X) = 6 = np \Rightarrow n = \frac{6}{p} = \frac{6}{0.75} = 8,$$

i.e., $X \sim \text{Binomial}(n = 8, p = 0.75)$.

(b) Qual o número modal de sensores que identificam corretamente uma peça defeituosa em n sensores?

(1.5)

Moda de X : número modal de sensores que identificam corretamente uma peça defeituosa em n sensores, denotada por m_0 .

Se $X \sim \text{Binomial}(n, p)$, então $P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$, $x = 0, 1, \dots, n$.

Sabe-se que m_0 é o valor de X para o qual a sua função massa de probabilidade é máxima, i.e., $m_0 \in \{0, 1, \dots, n\}$ satisfazendo i) $P(X = m_0) \geq P(X = m_0 - 1)$, ii) $P(X = m_0) \geq P(X = m_0 + 1)$.

Analisando a primeira desigualdade:

$$\frac{P(X = m_0)}{P(X = m_0 - 1)} = \frac{\binom{n}{m_0}(p)^{m_0}(1-p)^{n-m_0}}{\binom{n}{m_0-1}(p)^{m_0-1}(1-p)^{n-(m_0-1)}} = \frac{n-m_0+1}{m_0} \cdot \frac{p}{1-p} \geq 1 \quad \Leftrightarrow \quad m_0 \leq np + p$$

Analisando a segunda desigualdade:

$$\frac{P(X = m_0 + 1)}{P(X = m_0)} = \frac{\binom{n}{m_0+1}(p)^{m_0+1}(1-p)^{n-(m_0+1)}}{\binom{n}{m_0}(p)^{m_0}(1-p)^{n-m_0}} = \frac{n-m_0}{m_0+1} \cdot \frac{p}{1-p} \leq 1 \quad \Leftrightarrow \quad m_0 \geq np + p - 1$$

Para $X \sim \text{Binomial}(n = 8, p = 0.75)$

$$np + p - 1 = 5.75 \leq m_0 \leq np + p = 6.75 \quad \Rightarrow \quad m_0 = 6$$

Pergunta 3

2 valores

Num centro de atendimento, os clientes chegam segundo um processo em que o tempo entre chegadas consecutivas pode ser modelado por uma variável aleatória exponencial com média de 5 minutos. Um gestor pretende estudar o risco de o sistema permanecer inativo durante um período prolongado.

Seja X o tempo de espera até à próxima chegada de um cliente. Sabendo que já decorreram 3 minutos sem que tenha chegado qualquer cliente, qual é a probabilidade de ainda faltarem mais de 8 minutos para a chegada do próximo cliente?

A: 0.1108 B: 0.3679 C: 0.2018 D: 0.5488

X : tempo de espera (em minutos) até à próxima chegada de um cliente.

Como $X \sim \text{Exponencial}(\lambda)$ com $E(X) = \frac{1}{\lambda} = 5$, então $\lambda = \frac{1}{5}$.

Cálculo da probabilidade condicional $P(X > 11 \mid X > 3)$:

$$P(X > 11 \mid X > 3) = \frac{P(X > 11, X > 3)}{P(X > 3)} = \frac{P(X > 11)}{P(X > 3)} = \frac{1 - F_X(11)}{1 - F_X(3)} = \frac{e^{-11/5}}{e^{-3/5}} = \frac{0.1108}{0.5488} \approx 0.2018,$$

onde $F_X(x) = \int_0^x \lambda e^{-\lambda t} dt = e^{-\lambda t} \Big|_0^x = 1 - e^{-\lambda x} = 1 - e^{-x/5}$.

Alternativamente, pela propriedade de falta de memória da distribuição exponencial,

$$P(X > 11 \mid X > 3) = P(X > 8) = 1 - F_X(8) = e^{-8/5} \approx 0.2018$$

Pergunta 4

3 valores

Considere as variáveis aleatórias X e Y com função de probabilidade conjunta

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{3} \binom{x}{y} \left(\frac{1}{2}\right)^x, & x \in \{1, 2, 3\} \text{ e } y \in \{0, \dots, x\} \\ 0, & \text{caso contrário} \end{cases}$$

Tendo em conta que $\sum_{i=0}^k \binom{k}{i} = 2^k$ para $k \in \mathbb{N}^0$, obtenha o valor esperado de $2Y + 1$ condicional a $X = 3$.

Função de probabilidade marginal de $X = 3$:

$$f_X(3) = \sum_{y=0}^3 f_{X,Y}(3, y) = \sum_{y=0}^3 \frac{1}{3} \binom{3}{y} \left(\frac{1}{2}\right)^3 = \frac{1}{3} \left(\frac{1}{2}\right)^3 \sum_{y=0}^3 \binom{3}{y} = \frac{1}{3} \left(\frac{1}{2}\right)^3 \times 2^3 = \frac{1}{3}$$

Função de probabilidade condicional de $Y|X = 3$:

$$f_{Y|X=3}(y) = \frac{f_{X,Y}(3, y)}{f_X(3)} = 3f_{X,Y}(3, y), \text{ para } y \in \{0, \dots, 3\}$$

Valor esperado de $Y|X = 3$:

$$E(Y | X = 3) = \sum_{y=0}^3 y f_{Y|X=3}(y) = 3 \sum_{y=0}^3 y f_{X,Y}(3, y) = \sum_{y=0}^3 y \binom{3}{y} \left(\frac{1}{2}\right)^3 = 3 \times \frac{1}{2} = \frac{3}{2},$$

uma vez que $\sum_{y=0}^3 y \binom{3}{y} \left(\frac{1}{2}\right)^3$ é a expressão do valor esperado de uma variável aleatória com distribuição Binomial($n = 3, p = 1/2$).

Valor esperado condicional de $2Y + 1 | X = 3$:

$$E(2Y + 1 | X = 3) = 2E(Y | X = 3) + 1 = 2 \times \frac{3}{2} + 1 = 4$$

Pergunta 5

2 valores

Para garantir a viabilidade económica de um certo processo bioquímico, a concentração de oxigénio dissolvido (X) deve ser mantida num nível crítico. Sabe-se que X segue uma distribuição com valor esperado μ e variância σ^2 , ambas desconhecidas. Dois estagiários de engenharia, o João e a Maria, propõem estimadores diferentes para avaliar a concentração média, μ , com base numa amostra aleatória (X_1, X_2, X_3) de 3 medições independentes.

Estimador do João: $\hat{\mu}_J = \frac{X_1 + X_2 + X_3}{3}$

Estimador da Maria: $\hat{\mu}_M = \frac{X_1 + 2X_2 + X_3}{4}$

Considere as afirmações seguintes (**A1** a **A8**):

- | | |
|--|---|
| A1 $\hat{\mu}_J$ é um estimador centrado de μ . | A5 A variância de $\hat{\mu}_J$ é maior que a variância de $\hat{\mu}_M$ |
| A2 $\hat{\mu}_J$ é um estimador enviesado de μ | A6 A variância de $\hat{\mu}_M$ é maior que a variância de $\hat{\mu}_J$ |
| A3 $\hat{\mu}_M$ é um estimador centrado de μ | A7 O estimador $\hat{\mu}_J$ é mais eficiente do que o estimador $\hat{\mu}_M$ |
| A4 $\hat{\mu}_M$ é um estimador enviesado de μ | A8 O estimador $\hat{\mu}_M$ é mais eficiente do que o estimador $\hat{\mu}_J$ |

Selecione as afirmações correctas.

A: A1, A4, A6, A7 B: A1, A3, A6, A8 C: A1, A3, A5, A7 D: A1, A3, A6, A7

Usando as propriedades do valor esperado e da variância de combinações lineares de variáveis aleatórias, e o facto de X_i , $i = 1, 2, 3$ serem variáveis aleatórias independentes e identicamente distribuídas, com $E(X_i) = \mu$ e $V(X_i) = \sigma^2$, vem:

$$E(\hat{\mu}_J) = E\left(\frac{X_1 + X_2 + X_3}{3}\right) = \frac{E(X_1) + E(X_2) + E(X_3)}{3} = \frac{\mu + \mu + \mu}{3} = \mu$$

(logo $\hat{\mu}_J$ é centrado, ou não enviesado, pelo que **A1** está correcta e **A2** não está)

$$E(\hat{\mu}_M) = E\left(\frac{X_1 + 2X_2 + X_3}{4}\right) = \frac{E(X_1) + 2E(X_2) + E(X_3)}{4} = \frac{\mu + 2\mu + \mu}{4} = \mu$$

(logo $\hat{\mu}_M$ é centrado, ou não enviesado, pelo que **A3** está correcta e **A4** não está)

$$V(\hat{\mu}_J) = V\left(\frac{X_1 + X_2 + X_3}{3}\right) = \frac{V(X_1) + V(X_2) + V(X_3)}{9} = \frac{\sigma^2 + \sigma^2 + \sigma^2}{9} = \frac{\sigma^2}{3}$$

$$V(\hat{\mu}_M) = V\left(\frac{X_1 + 2X_2 + X_3}{4}\right) = \frac{V(X_1) + 4V(X_2) + V(X_3)}{16} = \frac{\sigma^2 + 4\sigma^2 + \sigma^2}{16} = \frac{3\sigma^2}{8} > \frac{\sigma^2}{3}$$

o que mostra que **A5** está incorrecta e **A6** está correcta. Por outro lado, um estimador é tanto mais eficiente quanto menor for a sua variância, pelo que **A7** está correcta e **A8** está incorrecta.

Num estudo sobre o tráfego de dados num *router* de uma rede empresarial, admite-se que o tempo (em microssegundo, μs) entre a chegada de pacotes sucessivos segue uma distribuição Exponencial de parâmetro $\lambda > 0$ desconhecido. Para uma concretização de uma amostra aleatória de X com dimensão $n = 225$, a soma total dos tempos entre chegadas foi de $90\mu s$.

- (a) Determine um intervalo de confiança aproximado a 98% para a taxa média de chegada de pacotes, λ (em pacotes/ μs). (2.0)

- X = tempo entre as chegadas de pacotes sucessivos (em μs)
 $X \sim \text{Exponencial}(\lambda)$, com $E(X) = 1/\lambda$ e $Var(X) = 1/\lambda^2$, λ desconhecido, $n = 225 \geq 30$
- Obtenção de IC aproximado para λ :

1. Seleção da v.a. fulcral para λ

Pelo Teorema do limite central: $Z = \frac{\bar{X} - 1/\lambda}{\sqrt{1/\lambda^2/n}} = (\lambda \bar{X} - 1)\sqrt{n} \stackrel{a}{\sim} N(0, 1)$

2. Obtenção dos quantis de probabilidade

Dado que $(1 - \alpha) = 0.98$ (i.e., $\alpha = 0.02$), usaremos os quantis simétricos:

$$b_\alpha = \Phi^{-1}(1 - \alpha/2) = \Phi^{-1}(0.99) \approx 2.3263$$

3. Inversão da desigualdade $-b_\alpha \leq Z \leq b_\alpha$

$$P[-b_\alpha \leq (\lambda \bar{X} - 1)\sqrt{n} \leq b_\alpha] \approx 1 - \alpha$$

$$\Leftrightarrow P\left[1 - \frac{b_\alpha}{\sqrt{n}} \leq \lambda \bar{X} \leq 1 + \frac{b_\alpha}{\sqrt{n}}\right] \approx 1 - \alpha$$

$$\Leftrightarrow P\left[\frac{1 - b_\alpha/\sqrt{n}}{\bar{X}} \leq \lambda \leq \frac{1 + b_\alpha/\sqrt{n}}{\bar{X}}\right] \approx 1 - \alpha$$

4. IC aproximado a $(1 - \alpha) \times 100\%$ para λ

$$IC(\lambda; (1 - \alpha) \times 100\%) \approx \left[\frac{1 - b_\alpha/\sqrt{n}}{\bar{x}}, \frac{1 + b_\alpha/\sqrt{n}}{\bar{x}}\right]$$

Dado que $n = 225$, $b_\alpha = 2.3263$ e $\bar{x} = \frac{90}{225} = 0.4$, temos:

$$IC(\lambda; 98\%) \approx \left[\frac{1 - 2.3263/15}{0.4}, \frac{1 + 2.3263/15}{0.4}\right] \approx [2.112, 2.888].$$

- (b) No âmbito do mesmo estudo, um relatório gerado por um *software* de monitorização apresentou, para os mesmos dados, um intervalo de confiança para λ com uma amplitude de 0.5 pacotes/ μs . Determine o nível de confiança aproximado configurado nesse *software*. (2.0)

- Obtenção do nível de confiança aproximado:

1. Identificação da expressão da amplitude (A)

A amplitude do intervalo de confiança para λ é a diferença entre os seus limites (obtida no Passo 3 da alínea anterior):

$$A = \frac{1 + b_\alpha/\sqrt{n}}{\bar{x}} - \frac{1 - b_\alpha/\sqrt{n}}{\bar{x}} = \frac{2b_\alpha}{\bar{x}\sqrt{n}}$$

2. Cálculo do quantil de probabilidade (b_α)

Substituindo os valores conhecidos ($A = 0.5$, $n = 225 \Rightarrow \sqrt{n} = 15$ e $\bar{x} = 0.4$):

$$0.5 = \frac{2b_\alpha}{0.4 \times 15} \Leftrightarrow 0.5 = \frac{2b_\alpha}{6} \Leftrightarrow b_\alpha = \frac{0.5 \times 6}{2} = 1.5$$

3. Determinação da probabilidade associada

O quantil $b_\alpha = 1.5$ relaciona-se com o nível de confiança através de $\Phi(b_\alpha) = 1 - \alpha/2$.

Consultando a tabela da distribuição Normal padrão:

$$\Phi(1.5) = P(Z \leq 1.5) \approx 0.9332$$

4. Cálculo do nível de confiança $(1 - \alpha) \times 100\%$

Sendo $1 - \alpha/2 = 0.9332$, temos:

$$\alpha/2 = 1 - 0.9332 = 0.0668 \Rightarrow \alpha = 0.1336$$

$$1 - \alpha = 1 - 0.1336 = 0.8664$$

O nível de confiança configurado no *software* é de aproximadamente 86.64%.

Pergunta 7

2 valores

Num estudo feito a um grupo de 13 nadadores, relativamente ao número de metros que cada um nada por semana observou-se um desvio padrão amostral $s = 7162$.

Com base nestes dados teste as hipóteses $H_0 : \sigma = 7000$ versus $H_1 : \sigma > 7000$, considerando um nível de significância de 10%. Indique os pressupostos necessários para validar o procedimento adotado.

- X : distância percorrida semanalmente, com pressuposto necessário $X \sim \text{Normal}(\mu, \sigma^2)$.
- $(X_1, \dots, X_n)$ amostra aleatória de X com concretização $n = 13$ e $s = 7162$ e distribuição $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$.
- Teste de hipóteses:
 1. Hipóteses: $H_0 : \sigma = 7000$ versus $H_1 : \sigma > 7000$.
 2. Estatística do teste: $T = \frac{(n-1)S^2}{7000^2} \underset{H_0}{\sim} \chi^2_{(12)}$, com valor observado $t = \frac{12 \times 7162^2}{7000^2} \approx 12.56186$.
 3. Região crítica (RC) a $\alpha = 0.1$: $RC_{10\%} = (c, +\infty)$, sendo $c = F_{\chi^2_{(12)}}^{-1}(0.9) = 18.54935$.
 4. Conclusão: Como $t = 12.56186 \notin RC_{10\%}$, não se deve rejeitar H_0 a 10%.

Pergunta 8

2 valores

Um engenheiro mecânico está a calibrar um sensor de momentos para um braço robótico industrial. O objetivo é verificar se o sensor apresenta algum erro sistemático, ou seja, se indica um valor diferente de zero quando nenhum momento é aplicado. Foram aplicados 12 níveis de momento conhecidos (x , em N.m) e registradas as leituras do sensor (Y , em N.m). Assume-se que o modelo de regressão linear simples, $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, é apropriado. Obtiveram-se os seguintes dados:

$$n = 12, \quad \bar{x} = 50, \quad \bar{y} = 50.8, \quad \sum_{i=1}^{12} (x_i - \bar{x})^2 = 2400, \quad \hat{\beta}_1 = 1.015, \quad \sum_{i=1}^{12} (y_i - \hat{y}_i)^2 = 2.50$$

Construa um intervalo de confiança a 95% para o parâmetro β_0 e use-o para concluir sobre a existência, ou não, de erro sistemático.

A: $IC(\beta_0, 0.95) = (-1.132, 1.232)$ e não existe erro sistemático

B: $IC(\beta_0, 0.95) = (-1.132, 1.232)$ e existe erro sistemático

C: $IC(\beta_0, 0.95) = (1.132, 2.232)$ e existe erro sistemático

- Modelo e hipóteses de trabalho: $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$, $i = 1, \dots, 12$, com $\varepsilon_i \stackrel{iid}{\sim} N(0, \sigma^2)$
- Método da variável fulcral (estimação intervalar de β_0):

1. Variável fulcral:

$$T = \frac{\hat{\beta}_0 - \beta_0}{\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right) \hat{\sigma}^2}} \sim t_{(n-2)} = t_{(10)}$$

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

$$P(a < T < b) = 1 - \alpha = 0.95 \Rightarrow a = -b \text{ e } b = F_{t_{(10)}}^{-1}(0.975) = 2.228$$

3. Desigualdade equivalente:

$$P\left(-b < \frac{\hat{\beta}_0 - \beta_0}{\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right) \hat{\sigma}^2}} < b\right) = 0.95$$

$$\Leftrightarrow P\left(\hat{\beta}_0 - b\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right) \hat{\sigma}^2} < \beta_0 < \hat{\beta}_0 + b\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right) \hat{\sigma}^2}\right) = 0.95,$$

4. Intervalo aleatório de confiança (IAC) e sua concretização (IC) a 95% para β_0 :

$$IAC(\beta_0; 0.95) : \hat{\beta}_0 \pm b\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right) \hat{\sigma}^2}$$

$$IC(\beta_0; 0.95) : 0.05 \pm 2.228 \times 0.5303 \approx (-1.132, 1.232),$$

tendo em conta os cálculos auxiliares abaixo.

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \Leftrightarrow \hat{\beta}_0 = 50.8 - (1.015 \times 50) = 0.05$$

$$\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right) \hat{\sigma}^2} = \sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (x_i - \bar{x})^2}\right) \hat{\sigma}^2} = \sqrt{\left(\frac{1}{12} + \frac{50^2}{2400}\right) \frac{2.5}{10}} \approx 0.5303$$

- Conclusão: Uma vez que o $IC(\beta_0; 0.95)$ contém o valor zero, é razoável concluir que o sensor não apresenta erro sistemático significativo.