

Duração: **120** minutos

Exame de Época Normal - 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

Num dado equipamento constituído por 12 transístores do mesmo tipo, em que 9 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, $i = 1, 2$.

$$P(F_1|\bar{F}_2) = \frac{P(F_1) \times P(\bar{F}_2|F_1)}{P(F_1) \times P(\bar{F}_2|F_1) + P(\bar{F}_1) \times P(\bar{F}_2|\bar{F}_1)} = \frac{\frac{9}{12} \times \frac{3}{11}}{\frac{9}{12} \times \frac{3}{11} + \frac{3}{12} \times \frac{2}{11}} = \frac{9}{11}$$

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:

- $E(X) = 8$
- $Var(X) = 1.6$

(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 $E(X) = np = 8$ e $Var(X) = np(1 - p) = 1.6$, pelo que

$$Var(X) = 1.6 = 8(1 - p) \Rightarrow p = 1 - \frac{1.6}{8} = 0.8 \quad \text{e} \quad E(X) = 8 = np \Rightarrow n = \frac{8}{p} = \frac{8}{0.8} = 10,$$

i.e., $X \sim \text{Binomial}(n = 10, p = 0.8)$.

(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 $X \sim \text{Binomial}(n = 10, p = 0.8)$

$$np + p - 1 = 7.8 \leq m_0 \leq np + p = 8.8 \quad \Rightarrow \quad m_0 = 8$$

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 4 minutos. Um gestor pretende estudar o risco de o sistema permanecer inactivo durante um período prolongado.

Seja X o tempo de espera até à próxima chegada de um cliente. Sabendo que já decorreram 2 minutos sem que tenha chegado qualquer cliente, qual é a probabilidade de ainda faltarem mais de 6 minutos para a chegada do próximo cliente?

A: 0.2231 B: 0.3679 C: 0.1353 D: 0.6065

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

Como $X \sim \text{Exponencial}(\lambda)$ com $E(X) = \frac{1}{\lambda} = 4$, então $\lambda = \frac{1}{4}$.

Cálculo da probabilidade condicional $P(X > 8 | X > 2)$:

$$P(X > 8 | X > 2) = \frac{P(X > 8, X > 2)}{P(X > 2)} = \frac{P(X > 8)}{P(X > 2)} = \frac{1 - F_X(8)}{1 - F_X(2)} = \frac{e^{-8/4}}{e^{-2/4}} = \frac{0.1353}{0.6065} \approx 0.2231,$$

onde $F_X(x) = \int_0^x \lambda e^{-\lambda t} dt = e^{-\lambda t} \Big|_0^x = 1 - e^{-\lambda x} = 1 - e^{-x/4}$.

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

$$P(X > 8 | X > 2) = P(X > 6) = 1 - F_X(6) = e^{-6/4} \approx 0.2231$$

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 $X = 2$.

Função de probabilidade marginal de $X = 2$:

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

Função de probabilidade condicional de $Y|X = 2$:

$$f_{Y|X=2}(y) = \frac{f_{X,Y}(2, y)}{f_X(2)} = 3f_{X,Y}(2, y), \text{ para } y \in \{0, \dots, 2\}$$

Valor esperado de $Y|X = 2$:

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

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

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

$$E(2Y + 1 | X = 2) = 2E(Y | X = 2) + 1 = 2 \times 1 + 1 = 3$$

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: $\hat{\mu}_M = \frac{X_1 + X_2 + 3X_3}{5}$

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

$$E(\hat{\mu}_M) = E\left(\frac{X_1 + X_2 + 3X_3}{5}\right) = \frac{E(X_1) + E(X_2) + 3E(X_3)}{5} = \frac{\mu + \mu + 3\mu}{5} = \mu$$

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

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

$$V(\hat{\mu}_M) = V\left(\frac{X_1 + X_2 + 3X_3}{5}\right) = \frac{V(X_1) + V(X_2) + 9V(X_3)}{25} = \frac{\sigma^2 + \sigma^2 + 9\sigma^2}{25} = \frac{11\sigma^2}{25} > \frac{\sigma^2}{3}$$

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 a eficiência de um servidor de base de dados, admite-se que o tempo (em microssegundo, μs) entre pedidos sucessivos segue uma distribuição Exponencial de parâmetro $\lambda > 0$ desconhecido. Para uma concretização de uma amostra aleatória de X com dimensão $n = 400$, a soma total dos tempos entre chegadas foi de $80\mu s$.

- (a) Determine um intervalo de confiança aproximado a 95% para a taxa média de chegada de pedidos, λ (em pedidos/ μs). (2.0)

- X = tempo entre as chegadas de pedidos sucessivos (em μs)
 $X \sim \text{Exponencial}(\lambda)$, com $E(X) = 1/\lambda$ e $\text{Var}(X) = 1/\lambda^2$, λ desconhecido, $n = 400 \geq 30$
- Obtenção de IC aproximado para λ

1. Seleção da v.a. fulcral para λ

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

2. Obtenção dos quantis de probabilidade

Para um nível de confiança $(1 - \alpha) = 0.95$, temos $\alpha = 0.05$.

Consultando a tabela da Normal padrão: $b_\alpha = \Phi^{-1}(0.975) = 1.96$

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

$$\begin{aligned} P[-b_\alpha \leq (\lambda \bar{X} - 1)\sqrt{n} \leq b_\alpha] &\simeq 1 - \alpha \\ \Leftrightarrow P\left[1 - \frac{b_\alpha}{\sqrt{n}} \leq \lambda \bar{X} \leq 1 + \frac{b_\alpha}{\sqrt{n}}\right] &\simeq 1 - \alpha \\ \Leftrightarrow P\left[\frac{1 - b_\alpha/\sqrt{n}}{\bar{X}} \leq \lambda \leq \frac{1 + b_\alpha/\sqrt{n}}{\bar{X}}\right] &\simeq 1 - \alpha \end{aligned}$$

4. Cálculo do IC a 95% para λ

Dados: $n = 400 \Rightarrow \sqrt{n} = 20$; $\bar{x} = \frac{80}{400} = 0.2$ ms.

$$IC(\lambda; 95\%) \simeq \left[\frac{1 - 1.96/20}{0.2}; \frac{1 + 1.96/20}{0.2} \right]$$

$$IC(\lambda; 95\%) \simeq \left[\frac{0.902}{0.2}; \frac{1.098}{0.2} \right] = [4.51; 5.49]$$

- (b) No âmbito do mesmo estudo, um relatório gerado por um *software* de diagnóstico apresentou, para os mesmos dados, um intervalo de confiança para λ com uma amplitude de 0.8 pedidos/ 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)

$$A = \frac{2b_\alpha}{\bar{x}\sqrt{n}}$$

2. Cálculo do quantil de probabilidade (b_α)

Substituindo os valores conhecidos ($A = 0.8$, $n = 400 \Rightarrow \sqrt{n} = 20$ e $\bar{x} = 0.2$):

$$0.8 = \frac{2b_\alpha}{0.2 \times 20} \Leftrightarrow 0.8 = \frac{2b_\alpha}{4} \Leftrightarrow b_\alpha = \frac{0.8 \times 4}{2} = 1.6$$

3. Determinação da probabilidade associada

Consultando a tabela da distribuição Normal padrão para $b_\alpha = 1.6$:

$$\Phi(1.6) = P(Z \leq 1.6) \approx 0.9452$$

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

Sendo $1 - \alpha/2 = 0.9452$, temos:

$$\alpha/2 = 1 - 0.9452 = 0.0548 \Rightarrow \alpha = 0.1096$$

$$1 - \alpha = 1 - 0.1096 = 0.8904$$

O nível de confiança configurado no *software* é de aproximadamente 89.04%.

Pergunta 7

2 valores

Num estudo feito a um grupo de 13 nadadoras, relativamente ao número de metros que cada uma nada por semana observou-se um desvio padrão amostral $s = 5700$.

Com base nestes dados teste as hipóteses $H_0 : \sigma = 6500$ versus $H_1 : \sigma < 6500$, considerando um nível de significância de 1%. 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 $s = 5700$ e distribuição $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$.
- Teste de hipóteses:
 1. Hipóteses: $H_0 : \sigma = 6500$ versus $H_1 : \sigma < 6500$.
 2. Estatística do teste: $T = \frac{(n-1)S^2}{6500^2} \underset{H_0}{\sim} \chi^2_{(12)}$, com valor observado $t = \frac{12 \times 5700^2}{6500^2} \approx 9.2279$.
 3. Região crítica (RC) a $\alpha = 0.01$: $RC_{1\%} = (0, c)$, sendo $c = F_{\chi^2_{(12)}}^{-1}(0.01) = 3.5706$.
 4. Conclusão: Como $t = 9.2279 \notin RC_{1\%}$, não se deve rejeitar H_0 a 1%.

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 99% para o parâmetro β_0 e use-o para concluir sobre a existência, ou não, de erro sistemático.

A: $IC(\beta_0, 0.99) = (-1.132, 1.232)$ e não existe erro sistemático

B: $IC(\beta_0, 0.99) = (-1.132, 1.232)$ e existe erro sistemático

C: $IC(\beta_0, 0.99) = (1.132, 2.232)$ e existe erro sistemático

D: $IC(\beta_0, 0.99) = (-1.631, 1.731)$ e não 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

$$P(a < T < b) = 1 - \alpha = 0.99 \Rightarrow a = -b \text{ e } b = F_{t_{(10)}}^{-1}(0.995) = 3.169$$

3. Desigualdade equivalente:

$$P\left(-b < \frac{\hat{\beta}_0 - \beta_0}{\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right) \hat{\sigma}^2}} < b\right) = 0.99$$

$$\Leftrightarrow P\left(\hat{\beta}_0 - b\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right) \hat{\sigma}^2} < \beta_0 < \hat{\beta}_0 + b\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right) \hat{\sigma}^2}\right) = 0.99,$$

4. Intervalo aleatório de confiança (IAC) e sua concretização (IC) a 99% para β_0 :

$$IAC(\beta_0; 0.99) : \hat{\beta}_0 \pm b\sqrt{\left(\frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n x_i^2 - n\bar{x}^2}\right) \hat{\sigma}^2}$$

$$IC(\beta_0; 0.99) : 0.05 \pm 3.169 \times 0.5303 \approx (-1.631, 1.731),$$

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 $IC(\beta_0; 0.99)$ contém o valor zero, é razoável concluir que o sensor não apresenta erro sistemático significativo.