

Duração: 120 minutos

Exame de Época Normal - Versão B

- 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

De uma embalagem de 10 componentes do mesmo tipo, em que 7 delas estão em perfeitas condições, são retiradas ao acaso, sucessivamente e sem reposição, 3 componentes.

- (a) Determine a probabilidade de todas as 3 componentes retiradas estarem em perfeitas condições. (1.5)

Acontecimentos e probabilidades:

$F_i = i$ -ésima componente retirada está em perfeitas condições, $i = 1, 2, 3$.

$$P(F_1 \cap F_2 \cap F_3) = P(F_1) \times P(F_2|F_1) \times P(F_3|F_1 \cap F_2) = \frac{7}{10} \times \frac{6}{9} \times \frac{5}{8} = \frac{7}{24}$$

- (b) Determine a probabilidade de a segunda componente retirada não estar em perfeitas condições. (1.5)

$$P(\bar{F}_2) = P(F_1 \cap \bar{F}_2) + P(\bar{F}_1 \cap \bar{F}_2) = P(F_1) \times P(\bar{F}_2|F_1) + P(\bar{F}_1) \times P(\bar{F}_2|\bar{F}_1) = \frac{7}{10} \times \frac{3}{9} + \frac{3}{10} \times \frac{2}{9} = \frac{3}{10}$$

Pergunta 2

2 valores

Num sistema de transmissão de dados, são enviados 500 *bits* através de um canal ruidoso. Cada *bit* é recebido incorretamente (erro de transmissão) com probabilidade 0.002, independentemente dos restantes.

- Seja X o número de erros de transmissão em 500 *bits* enviados.
- O custo associado ao processo é dado por $C(X) = X^2$.

Determine o custo esperado do processo.

A: 1 B: 1.998 C: 0.998 D: 0.002

X : número de erros de transmissão em 500 *bits* enviados.

Como cada *bit* pode ser transmitido incorretamente com probabilidade $p = 0.002$, independentemente dos restantes, têm-se $n = 500$ ensaios de Bernoulli ($p = 0.002$) independentes $\Rightarrow X \sim \text{Binomial}(500, 0.002)$. Valor esperado do processo:

$$E[C(X)] = E(X^2) = \text{Var}(X) + (E(X))^2 = np(1-p) + (np)^2 = 500 \times 0.002 \times 0.998 + (500 \times 0.002)^2 = 1.998$$

Pergunta 3

3 valores

Um sistema de detecção de intrusão analisa pacotes de dados em blocos de n unidades. Seja Y a variável aleatória que representa o número de pacotes infectados num bloco, seguindo uma distribuição binomial, i.e., $Y \sim \text{Binomial}(n, p)$.

- (a) Usando a definição de valor esperado de uma variável aleatória, mostre que $E(Y) = np$. (1.0)

Se $Y \sim \text{Binomial}(n, p)$, a sua função massa de probabilidade é $f_Y(y) = \binom{n}{y} p^y (1-p)^{n-y}$, $y = 0, 1, \dots, n$, tal que $\sum_{y=0}^n f_Y(y) = \sum_{y=0}^n \binom{n}{y} p^y (1-p)^{n-y} = 1$ (propriedade P1), pelo que

$$\begin{aligned} E(Y) &= \sum_{y=0}^n y f_Y(y) = \sum_{y=0}^n y \binom{n}{y} p^y (1-p)^{n-y} = \sum_{y=1}^n \frac{y n (n-1)!}{y (y-1)! (n-y)!} p^{y+1-1} (1-p)^{n-y} \\ &= n p \sum_{y=1}^{n-1} \binom{n-1}{y-1} p^{y-1} (1-p)^{n-1-(y-1)} \stackrel{z=y-1}{=} n p \sum_{z=0}^{n-1} \binom{n-1}{z} p^z (1-p)^{n-1-z} \\ &\stackrel{P1}{=} n p \times 1 = n p \end{aligned}$$

- (b) Sabe-se que o valor esperado de pacotes infectados por bloco é $E(Y) = 0.8$ e a variância é $\text{Var}(Y) = 0.72$. (2.0)
Determine os valores de n e p e a probabilidade de um bloco conter exatamente 3 pacotes infectados, sabendo que o bloco não está limpo (ou seja, contém pelo menos um pacote infectado).

Determinação dos parâmetros n e p :

Sabe-se que $E(Y) = np = 0.8$ e $\text{Var}(Y) = np(1-p) = 0.72$, pelo que

$$\text{Var}(Y) = 0.72 = 0.8(1-p) \Rightarrow p = 1 - \frac{0.72}{0.8} = 0.1 \quad \text{e} \quad E(Y) = 0.8 = np \Rightarrow n = \frac{0.8}{p} = \frac{0.8}{0.1} = 8,$$

i.e., $Y \sim \text{Binomial}(n = 8, p = 0.1)$.

Cálculo da probabilidade condicional $P(Y = 3 | Y \geq 1)$:

$$P(Y = 3 | Y \geq 1) = \frac{P(Y = 3, Y \geq 1)}{P(Y \geq 1)} = \frac{P(Y = 3)}{P(Y \geq 1)} = \frac{F_Y(3) - F_Y(2)}{1 - F_Y(0)} = \frac{0.9950 - 0.9619}{1 - 0.4305} = \frac{0.0331}{0.5695} \approx 0.0581$$

Pergunta 4

2 valores

Suponha que numa região sísmica os grandes sismos ocorrem ao longo do tempo segundo uma distribuição de Poisson ($\lambda = 2$ por ano) e que essas ocorrências são independentes em períodos não sobrepostos. Calcule a probabilidade (aproximada) de em 50 anos a média das ocorrências de grandes sismos ocorrerem nem mais de 2.4 nem menos de 1.5.

A: 0.9772 B: 1 C: 0.9710 D: 0.0062

- X_i : número de grandes sismos na região sísmica no ano i , $i = 1, \dots, 50$.
- $\bar{X} = \frac{1}{50} \sum_{i=1}^{50} X_i$ é a média das ocorrências de grandes sismos em 50 anos.
- Se $X_i \sim \text{Poisson}(\lambda = 2)$, $E(X_i) = 2 = \text{Var}(X_i)$, $\forall i$ (i.d.) e $X_1, \dots, X_{50}$ independentes (ind.),

$$\begin{aligned} E(\bar{X}) &= \frac{1}{50} E\left(\sum_{i=1}^{50} X_i\right) = \frac{1}{50} \sum_{i=1}^{50} E(X_i) \stackrel{i.d.}{=} \frac{1}{50} \times 50 \times 2 = 2 \\ \text{Var}(\bar{X}) &= \left(\frac{1}{50}\right)^2 \text{Var}\left(\sum_{i=1}^{50} X_i\right) \stackrel{ind.}{=} \left(\frac{1}{50}\right)^2 \sum_{i=1}^{50} \text{Var}(X_i) \stackrel{i.d.}{=} \left(\frac{1}{50}\right)^2 \times 50 \times 2 = \frac{2}{50} \end{aligned}$$

- Para grandes amostras ($n = 50$), pelo T.L.C., $\bar{X} \stackrel{a}{\sim} N(2, 2/50)$, i.e., $Z = \frac{\bar{X}-2}{\sqrt{2/50}} \stackrel{a}{\sim} N(0, 1)$.

$$\begin{aligned} P(1.5 \leq \bar{X} \leq 2.4) &\approx P\left(\frac{1.5-2}{\sqrt{2/50}} \leq Z \leq \frac{2.4-2}{\sqrt{2/50}}\right) = P(-2.5 \leq Z \leq 2) = F_{N(0,1)}(2) - F_{N(0,1)}(-2.5) \\ &= 0.9772 - 0.0062 \approx 0.9710 \end{aligned}$$

Um engenheiro informático está a analisar o tempo de resposta de um novo serviço de autenticação, tendo concluído que esse tempo, X (em milésimos de segundo), é uma variável aleatória contínua, com a seguinte função densidade de probabilidade,

$$f(x; \lambda) = \lambda e^{-\lambda(x-1)}, \quad x \geq 1, \quad \lambda > 0.$$

Selecioneados aleatoriamente 5 pedidos de autenticação, observaram-se os seguintes tempos de resposta: (1.1252, 1.0524, 1.0598, 1.1359, 1.0624). Use essa amostra para calcular uma estimativa de máxima verosimilhança de λ e da probabilidade de X exceder 1.2 milésimos de segundo.

- Seja $(X_1, \dots, X_n)$ uma amostra aleatória de X .
- Método da máxima verosimilhança (estimação de λ):

1. Funções de verosimilhança e de log-verosimilhança:

$$L(\lambda|\mathbf{x}) = f_{\mathbf{X}}(\mathbf{x}|\lambda) \stackrel{i.i.d.}{=} \prod_{i=1}^n f_X(x_i|\lambda) = \prod_{i=1}^n \lambda e^{-\lambda(x_i-1)} = \lambda^n e^{-\lambda \sum_{i=1}^n (x_i-1)}$$

$$\ln L(\lambda|\mathbf{x}) = n \ln \lambda - \lambda \left(\sum_{i=1}^n (x_i - 1) \right)$$

2. Maximização de $\ln L(\lambda|\mathbf{x})$, com $\hat{\lambda}$ a denotar a estimativa de MV de λ :

$$\left. \frac{d}{d\lambda} \ln L(\lambda|\mathbf{x}) \right|_{\lambda=\hat{\lambda}} = 0 \Leftrightarrow \frac{n}{\lambda} - \sum_{i=1}^n (x_i - 1) = 0 \Rightarrow \hat{\lambda} = \frac{n}{\sum_{i=1}^n (x_i - 1)} = \frac{1}{\bar{x} - 1}$$

$$\left. \frac{d^2}{d\lambda^2} \ln L(\lambda|\mathbf{x}) \right|_{\lambda=\hat{\lambda}} < 0 \Rightarrow -\frac{n}{\hat{\lambda}^2} < 0 \quad (\hat{\lambda} \text{ é ponto de máximo})$$

- Concretização, estimativa de máxima verosimilhança de λ :

Usando a amostra (1.1252, 1.0524, 1.0598, 1.1359, 1.0624), com $n = 5$ e $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = 1.0871$, conclui-se que a estimativa de máxima verosimilhança de λ é

$$\hat{\lambda} = \frac{1}{1.0871 - 1} \simeq 11.48.$$

- Pela propriedade da invariância dos estimadores de máxima verosimilhança, a estimativa de máxima verosimilhança de $P(X > 1.2) = \int_{1.2}^{\infty} \lambda e^{-\lambda(x-1)} dx = -e^{-\lambda(x-1)} \Big|_{1.2}^{\infty} = e^{-0.2\lambda}$ é

$$P(\widehat{X > 1.2}) = e^{-0.2\hat{\lambda}} = e^{-0.2 \times 11.48} = 0.1007$$

Admita que a variável aleatória X , que representa o comprimento de uma componente mecânica (em mm), pode ser modelada por uma distribuição Normal de média μ e variância σ^2 , sendo ambos os parâmetros desconhecidos. Com base na concretização de uma amostra aleatória $(X_1, \dots, X_{11})$ que resultou em:

$$\sum_{i=1}^{11} x_i = 110 \quad \text{e} \quad \sum_{i=1}^{11} x_i^2 = 1190,$$

indique um intervalo de confiança a 99% para o desvio padrão do comprimento de uma componente mecânica:

A: (3.573, 41.747) **B:** (1.890, 6.461) **C:** (1.969, 5.931) **D:** (2.802, 6.160)

- Seja $(X_1, \dots, X_n)$ uma amostra aleatória de $X \sim N(\mu, \sigma^2)$ e a sua concretização para $n = 11$ conduziu a: $\sum_{i=1}^{11} x_i = 110$ e $\sum_{i=1}^{11} x_i^2 = 1190$, implicando $s^2 = \frac{1}{11-1} \left(\sum_{i=1}^{11} x_i^2 - \frac{(\sum_{i=1}^{11} x_i)^2}{11} \right) = 9$.
- Método da variável fulcral (estimação intervalar de σ):

1. Variável fulcral: $W = \frac{(n-1)S^2}{\sigma^2} \sim \chi_{(n-1)}^2 = \chi_{(10)}^2$.

2. Fixado um grau de confiança $1 - \alpha = 0.99$, obtêm-se a e b tais que

$$P(a < W < b) = 1 - \alpha = 0.99 \Rightarrow a = F_{\chi_{(10)}^2}^{-1}(0.005) = 2.156 \quad \text{e} \quad b = F_{\chi_{(10)}^2}^{-1}(0.995) = 25.188$$

3. Desigualdade equivalente:

$$P\left(a < \frac{(n-1)S^2}{\sigma^2} < b\right) = 0.99 \Leftrightarrow P\left(\sqrt{\frac{(n-1)S^2}{b}} < \sigma < \sqrt{\frac{(n-1)S^2}{a}}\right) = 0.99$$

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

$$IAC(\sigma; 0.99): \left(\sqrt{\frac{(n-1)S^2}{b}}, \sqrt{\frac{(n-1)S^2}{a}} \right), \quad IC(\sigma; 0.99): \left(\sqrt{\frac{10 \times 9}{25.188}}, \sqrt{\frac{10 \times 9}{2.156}} \right) = (1.890, 6.461)$$

Pergunta 7

3 valores

Considere que a voltagem à saída num circuito elétrico segue uma distribuição normal com desvio padrão igual a 2.1. Em condições regulares espera-se desta voltagem um valor igual a 130.

- (a) Numa amostra independente e identicamente distribuída de 40 medições da voltagem neste circuito, obteve-se uma média igual a 129.6. Com base nesta amostra avalie se o circuito encontra-se a funcionar dentro da regularidade, considerando um nível de significância de 1%. (2.0)

- X : voltagem à saída de circuito elétrico, com $X \sim \text{Normal}(\mu, \sigma^2 = 2.1^2)$.
- $(X_1, \dots, X_n)$ amostra aleatória de X com concretização $n = 40$ e $\bar{x} = 129.6$ e distribuição amostral $\bar{X} \sim \text{Normal}(\mu, \sigma^2/n)$, sendo $\sigma^2/n = 2.1^2/40$.
- Teste de hipóteses:
 1. Hipóteses: $H_0: \mu = 130$ versus $H_1: \mu \neq 130$.
 2. Estatística do teste: $T = \frac{\bar{X} - 130}{2.1/\sqrt{40}} \underset{H_0}{\sim} N(0, 1)$, com valor observado $t = \frac{129.6 - 130}{2.1/\sqrt{40}} \approx -1.2047$.
 3. Região crítica (RC) a $\alpha = 0.01$: $RC_{1\%} = (-\infty, -c) \cup (c, +\infty)$, sendo $c = \Phi^{-1}(0.995) = 2.5758$.
 4. Conclusão: Como $t = -1.2047 \notin RC_{1\%}$, não se deve rejeitar H_0 a 1%, o que indica que o circuito se encontra a funcionar dentro da regularidade.

- (b) Tendo em conta o cenário da alínea anterior, calcule a probabilidade dos erros: (1.0)

- de 1ª espécie (rejeitar H_0 quando H_0 é verdadeiro),
- de 2ª espécie (aceitar H_0 quando H_0 é falso)

para o teste de hipóteses $H_0: \mu = 130$ versus $H_1: \mu = 129$.

Sejam α e β as probabilidades do erro de 1ª e 2ª espécie, respetivamente, com hipóteses $H_0 : \mu = 130 \equiv \mu_0$ e $H_1 : \mu = 129 \equiv \mu_1$.

$$\begin{aligned}\alpha &= P(\text{Rejeitar } H_0 | H_0 \text{ verdadeiro}) = P\left(\left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| \geq 2.5758 \mid X \sim N(130, 2.1^2)\right) \\ &= 1 - (\Phi(2.5758) - \Phi(-2.5758)) \approx 1 - 0.99 = 0.01\end{aligned}$$

$$\begin{aligned}\beta &= P(\text{Aceitar } H_0 | H_1 \text{ verdadeiro}) = P\left(\left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| \leq 2.5758 \mid X \sim N(129, 2.1^2)\right) \\ &= P\left(-2.5758 \frac{\sigma}{\sqrt{n}} + \mu_0 \leq \bar{X} \leq 2.5758 \frac{\sigma}{\sqrt{n}} + \mu_0 \mid X \sim N(129, 2.1^2)\right) \\ &= P\left(\frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} - 2.5758 \leq \frac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} \leq \frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} + 2.5758 \mid X \sim N(129, 2.1^2)\right) \\ &= \Phi(5.5875) - \Phi(0.4359) \approx 1 - 0.6686 = 0.3314\end{aligned}$$

Alternativa (teste unilateral): Para a região crítica $RC_{1\%} = (-\infty, -b)$, onde $b = \Phi^{-1}(0.99) = 2.33$,

$$\begin{aligned}\alpha &= P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \in RC_{1\%} \mid \mu = \mu_0 \equiv 130\right) = P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < -2.33 \mid \mu = 130\right) = \Phi(-2.33) = 0.01 \\ \beta &= P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \notin RC_{1\%} \mid \mu = \mu_1 \equiv 129\right) = P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > -2.33 \mid \mu = 129\right) \\ &= P\left(\bar{X} > -2.33 \frac{\sigma}{\sqrt{n}} + 130 \mid \mu = 129\right) = P\left(\frac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} > -2.33 + \frac{130 - 129}{\sigma/\sqrt{n}} \mid \mu = 129\right) \\ &= 1 - \Phi(0.6817) = 1 - 0.7523 = 0.2477\end{aligned}$$

Pergunta 8

2 valores

No controlo de qualidade de uma linha de produção, estuda-se o tempo de secagem de uma componente eletrónica, X , em horas. Admita que X é uma variável aleatória contínua com a seguinte função densidade de probabilidade:

$$f_X(x) = \frac{2}{9}x, \quad 0 < x < 3.$$

Uma amostra de $n = 100$ componentes foi recolhida e os tempos observados foram agrupados na tabela abaixo:

Tempo de secagem (horas)	]0, 1.5[	[1.5, 2.4[	[2.4, 3[
Frequência absoluta observada (o_i)	22	42	36

Averigue o ajustamento dos dados a este modelo teórico, indicando o valor-p aproximado do teste.

A: 0.0035 **B:** 0.1015 **C:** 0.7442 **D:** 0.8984

- Como $f_X(x) = \frac{2}{9}x$, para $0 < x < 3$, então a função de distribuição de X é:

$$F_X(x) = \int_0^x f_X(u) du = \int_0^x \frac{2}{9}u du = \left[\frac{u^2}{9} \right]_0^x = \frac{x^2}{9}, 0 < x < 3$$

- Teste de ajustamento do Qui-quadrado de Pearson com $n = 100$ e $k = 3$ classes.

1. Hipóteses: $H_0 : f_X(x) = \frac{2}{9}x$ versus $H_1 : f_X(x) \neq \frac{2}{9}x$, $0 < x < 3$.

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

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

- $p_1^0 = F_X(1.5) - F_X(0) = 1.5^2/9 = 0.25$
- $p_2^0 = F_X(2.4) - F_X(1.5) = 2.4^2/9 - 0.25 = 0.64 - 0.25 = 0.39$
- $p_3^0 = F_X(3) - F_X(2.4) = 1 - 0.64 = 0.36$

i	classe(x) _{i}	o_i	$E_i = 100 \times p_i^0$	$\frac{(o_i - E_i)^2}{E_i}$
1	]0, 1.5[	22	25	0.3600
2	[1.5, 2.4[	42	39	0.2308
3	[2.4, 3[	36	36	0.0000
total		100	100	$t \approx 0.5908$

3. Valor-p: $P = P(T \geq 0.5908 | H_0) \approx 1 - F_{\chi_{(2)}^2}(0.5908) = 1 - 0.2558 = 0.7442$. Note-se que $P \in (0.7, 0.8) = (1 - F_{\chi_{(2)}^2}(0.446), 1 - F_{\chi_{(2)}^2}(0.713))$.

4. Conclusão: Com valor-p = 0.7442, não se rejeita H_0 para os níveis de significância usuais ($\alpha = 1\%, 5\%$ ou 10%).