

Duração: 120 minutos

Exame de Época Normal - 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

De uma embalagem de 12 componentes do mesmo tipo, em que 8 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, $i = 1, 2, 3$.

$$P(F_1 \cap F_2 \cap F_3) = P(F_1) \times P(F_2|F_1) \times P(F_3|F_1 \cap F_2) = \frac{8}{12} \times \frac{7}{11} \times \frac{6}{10} = \frac{14}{55}$$

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

$$P(\bar{F}_2) = P(F_1 \cap \bar{F}_2) + P(\bar{F}_1 \cap \bar{F}_2) = P(F_1) \times P(\bar{F}_2|F_1) + P(\bar{F}_1) \times P(\bar{F}_2|\bar{F}_1) = \frac{8}{12} \times \frac{4}{11} + \frac{4}{12} \times \frac{3}{11} = \frac{1}{3}$$

Pergunta 2

2 valores

Num sistema de transmissão de dados, são enviados 800 *bits* através de um canal ruidoso. Cada *bit* é recebido incorretamente (erro de transmissão) com probabilidade 0.0015, independentemente dos restantes.

- Seja X o número de erros de transmissão em 800 *bits* enviados.
- O custo associado ao processo é dado por $C(X) = X^2$.

Determine o custo esperado do processo.

A: 1.44 B: 1.1982 C: 1.2 D: 2.6382

X: número de erros de transmissão em 800 *bits* enviados.

Como cada *bit* pode ser transmitido incorretamente com probabilidade $p = 0.0015$, independentemente dos restantes, têm-se $n = 800$ ensaios de Bernoulli($p = 0.0015$) independentes $\Rightarrow X \sim \text{Binomial}(800, 0.0015)$.

Valor esperado do processo:

$$E[C(X)] = E(X^2) = \text{Var}(X) + (E(X))^2 = np(1-p) + (np)^2 = 800 \times 0.0015 \times 0.9985 + (800 \times 0.0015)^2 = 2.6382$$

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 $Y \sim \text{Binomial}(n, p)$.

- (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

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

- (b) Sabe-se que o valor esperado de pacotes infectados por bloco é $E(Y) = 1.2$ e a variância é $\text{Var}(Y) = 0.96$. (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 $E(Y) = np = 1.2$ e $\text{Var}(Y) = np(1-p) = 0.96$, pelo que

$$\text{Var}(Y) = 0.96 = 1.2(1-p) \Rightarrow p = 1 - \frac{0.96}{1.2} = 0.2 \quad \text{e} \quad E(Y) = 1.2 = np \Rightarrow n = \frac{1.2}{p} = \frac{1.2}{0.2} = 6,$$

i.e., $Y \sim \text{Binomial}(n = 6, p = 0.2)$.

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

$$P(Y = 3 | Y \geq 1) = \frac{P(Y = 3, Y \geq 1)}{P(Y \geq 1)} = \frac{P(Y = 3)}{P(Y \geq 1)} = \frac{F_Y(3) - F_Y(2)}{1 - F_Y(0)} = \frac{0.9830 - 0.9011}{1 - 0.2621} = \frac{0.0819}{0.7379} \approx 0.111$$

Pergunta 4

2 valores

Suponha que numa região sísmica os grandes sismos ocorrem ao longo do tempo segundo uma distribuição de Poisson ($\lambda = 3$ por ano) e que essas ocorrências são independentes em períodos não sobrepostos. Calcule a probabilidade (aproximada) de em 60 anos a média das ocorrências de grandes sismos ocorrerem nem mais de 3.2 nem menos de 2.6.

A: 0.7776 B: 1 C: 0.8144 D: 0.0368

- X_i : número de grandes sismos na região sísmica no ano i , $i = 1, \dots, 60$.
- $\bar{X} = \frac{1}{60} \sum_{i=1}^{60} X_i$ é a média das ocorrências de grandes sismos em 60 anos.
- Se $X_i \sim \text{Poisson}(\lambda = 3)$, $E(X_i) = 3 = \text{Var}(X_i)$, $\forall i$ (i.d.) e $X_1, \dots, X_{60}$ independentes (ind.),

$$\begin{aligned} E(\bar{X}) &= \frac{1}{60} E\left(\sum_{i=1}^{60} X_i\right) = \frac{1}{60} \sum_{i=1}^{60} E(X_i) \stackrel{i.d.}{=} \frac{1}{60} \times 60 \times 3 = 3 \\ \text{Var}(\bar{X}) &= \left(\frac{1}{60}\right)^2 \text{Var}\left(\sum_{i=1}^{60} X_i\right) \stackrel{ind.}{=} \left(\frac{1}{60}\right)^2 \sum_{i=1}^{60} \text{Var}(X_i) \stackrel{i.d.}{=} \left(\frac{1}{60}\right)^2 \times 60 \times 3 = \frac{3}{60} \end{aligned}$$

- Para grandes amostras ($n = 60$), pelo T.L.C., $\bar{X} \stackrel{a}{\sim} N(3, 3/60)$, i.e., $Z = \frac{\bar{X} - 3}{\sqrt{3/60}} \stackrel{a}{\sim} N(0, 1)$.

$$\begin{aligned} P(2.6 \leq \bar{X} \leq 3.2) &\approx P\left(\frac{2.6-3}{\sqrt{3/60}} \leq Z \leq \frac{3.2-3}{\sqrt{3/60}}\right) = P(-1.7888 \leq Z \leq 0.8944) = F_{N(0,1)}(0.8944) - F_{N(0,1)}(-1.7888) \\ &= 0.8144 - 0.0368 \approx 0.7776 \end{aligned}$$

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,

$$f(x; \lambda) = \lambda e^{-\lambda(x-2)}, \quad x \geq 2, \quad \lambda > 0.$$

Selecionados aleatoriamente 4 pedidos de autenticação, observaram-se os seguintes tempos de resposta: (2.1252, 2.0524, 2.0598, 2.1359). Use essa amostra para calcular uma estimativa de máxima verosimilhança de λ e da probabilidade de X exceder 2.2 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:

$$L(\lambda|\mathbf{x}) = f_{\mathbf{X}}(\mathbf{x}|\lambda) \stackrel{i.i.d.}{=} \prod_{i=1}^n f_X(x_i|\lambda) = \prod_{i=1}^n \lambda e^{-\lambda(x_i-2)} = \lambda^n e^{-\lambda \sum_{i=1}^n (x_i-2)}$$

$$\ln L(\lambda|\mathbf{x}) = n \ln \lambda - \lambda \left(\sum_{i=1}^n (x_i - 2) \right)$$

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

$$\left. \frac{d}{d\lambda} \ln L(\lambda|\mathbf{x}) \right|_{\lambda=\hat{\lambda}} = 0 \Leftrightarrow \frac{n}{\lambda} - \sum_{i=1}^n (x_i - 2) = 0 \Rightarrow \hat{\lambda} = \frac{n}{\sum_{i=1}^n (x_i - 2)} = \frac{1}{\bar{x} - 2}$$

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

- Concretização, estimativa de máxima verosimilhança de λ :

Usando a amostra (2.1252, 2.0524, 2.0598, 2.1359), com $n = 4$ e $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = 2.0933$, conclui-se que a estimativa de máxima verosimilhança de λ é

$$\hat{\lambda} = \frac{1}{2.0933 - 2} \simeq 10.72.$$

- Pela propriedade da invariância dos estimadores de máxima verosimilhança, a estimativa de máxima verosimilhança de $P(X > 2.2) = \int_{2.2}^{\infty} \lambda e^{-\lambda(x-2)} dx = -e^{-\lambda(x-2)} \Big|_{2.2}^{\infty} = e^{-0.2\lambda}$ é

$$P(\widehat{X > 2.2}) = e^{-0.2\hat{\lambda}} = e^{-0.2 \times 10.72} = 0.1171$$

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 $(X_1, \dots, X_{16})$ que resultou em:

$$\sum_{i=1}^{16} x_i = 320 \quad \text{e} \quad \sum_{i=1}^{16} x_i^2 = 7000,$$

indique um intervalo de confiança a 95% para o desvio padrão do comprimento de uma componente mecânica:

A: (4.671, 9.789) **B:** (21.828, 95.816) **C:** (4.522, 10.477) **D:** (5.011, 8.528)

- Seja $(X_1, \dots, X_n)$ uma amostra aleatória de $X \sim N(\mu, \sigma^2)$ e a sua concretização para $n = 16$ conduziu a: $\sum_{i=1}^{16} x_i = 320$ e $\sum_{i=1}^{16} x_i^2 = 7000$, implicando $s^2 = \frac{1}{16-1} \left(\sum_{i=1}^{16} x_i^2 - \frac{(\sum_{i=1}^{16} x_i)^2}{16} \right) = 40$.
- Método da variável fulcral (estimação intervalar de σ):

1. Variável fulcral: $W = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)} = \chi^2_{(15)}$.

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

$$P(a < W < b) = 1 - \alpha = 0.95 \Rightarrow a = F_{\chi^2_{(15)}}^{-1}(0.025) = 6.262 \quad \text{e} \quad b = F_{\chi^2_{(15)}}^{-1}(0.975) = 27.488$$

3. Desigualdade equivalente:

$$P\left(a < \frac{(n-1)S^2}{\sigma^2} < b\right) = 0.95 \Leftrightarrow P\left(\sqrt{\frac{(n-1)S^2}{b}} < \sigma < \sqrt{\frac{(n-1)S^2}{a}}\right) = 0.95$$

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

$$IAC(\sigma; 0.95): \left(\sqrt{\frac{(n-1)S^2}{b}}, \sqrt{\frac{(n-1)S^2}{a}} \right), IC(\sigma; 0.95): \left(\sqrt{\frac{15 \times 40}{27.488}}, \sqrt{\frac{15 \times 40}{6.262}} \right) = (4.671, 9.789)$$

Pergunta 7

3 valores

Considere que a voltagem à saída num circuito elétrico segue uma distribuição normal com desvio padrão igual a 1.1. Em condições regulares espera-se desta voltagem um valor igual a 120.

- (a) Numa amostra independente e identicamente distribuída de 40 medições da voltagem neste circuito, obteve-se uma média igual a 120.4. Com base nesta amostra avalie se o circuito encontra-se a funcionar dentro da regularidade, considerando um nível de significância de 10%. (2.0)

- X : voltagem à saída de circuito elétrico, com $X \sim \text{Normal}(\mu, \sigma^2 = 1.1^2)$.
- $(X_1, \dots, X_n)$ amostra aleatória de X com concretização $n = 40$ e $\bar{x} = 120.4$ e distribuição amostral $\bar{X} \sim \text{Normal}(\mu, \sigma^2/n)$, sendo $\sigma^2/n = 1.1^2/40$.
- Teste de hipóteses:
 1. Hipóteses: $H_0: \mu = 120$ versus $H_1: \mu \neq 120$.
 2. Estatística do teste: $T = \frac{\bar{X} - 120}{1.1/\sqrt{40}} \underset{H_0}{\sim} N(0, 1)$, com valor observado $t = \frac{120.4 - 120}{1.1/\sqrt{40}} \approx 2.2998$.
 3. Região crítica (RC) a $\alpha = 0.1$: $RC_{10\%} = (-\infty, -c) \cup (c, +\infty)$, sendo $c = \Phi^{-1}(0.95) = 1.6448$.
 4. Conclusão: Como $t = 2.2998 \in RC_{10\%}$, deve-se rejeitar H_0 a 10%, o que indica que o circuito não 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 $H_0: \mu = 120$ versus $H_1: \mu = 120.5$.

Sejam α e β as probabilidades do erro de 1ª e 2ª espécie, respetivamente, com hipóteses $H_0 : \mu = 120 \equiv \mu_0$ e $H_1 : \mu = 120.5 \equiv \mu_1$.

$$\begin{aligned}\alpha &= P(\text{Rejeitar } H_0 | H_0 \text{ verdadeiro}) = P\left(\left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| \geq 1.6448 \mid X \sim N(120, 1.1^2)\right) \\ &= 1 - (\Phi(1.6448) - \Phi(-1.6448)) \approx 1 - 0.9 = 0.1\end{aligned}$$

$$\begin{aligned}\beta &= P(\text{Aceitar } H_0 | H_1 \text{ verdadeiro}) = P\left(\left|\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right| \leq 1.6448 \mid X \sim N(120.5, 1.1^2)\right) \\ &= P\left(-1.6448 \frac{\sigma}{\sqrt{n}} + \mu_0 \leq \bar{X} \leq 1.6448 \frac{\sigma}{\sqrt{n}} + \mu_0 \mid X \sim N(120.5, 1.1^2)\right) \\ &= P\left(\frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} - 1.6448 \leq \frac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} \leq \frac{\mu_0 - \mu_1}{\sigma/\sqrt{n}} + 1.6448 \mid X \sim N(120.5, 1.1^2)\right) \\ &= \Phi(-4.5196) - \Phi(-1.2299) \approx 0.1094 - 0 = 0.1094\end{aligned}$$

Alternativa (teste unilateral): Para a região crítica $RC_{10\%} = (b, \infty)$, onde $b = \Phi^{-1}(0.90) = 1.28$,

$$\begin{aligned}\alpha &= P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \in RC_{10\%} \mid \mu = \mu_0 \equiv 120\right) = P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} > 1.28 \mid \mu = 120\right) = 1 - \Phi(1.28) = 0.1 \\ \beta &= P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \notin RC_{10\%} \mid \mu = \mu_1 \equiv 120.5\right) = P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < 1.28 \mid \mu = 120.5\right) \\ &= P\left(\bar{X} < 1.28 \frac{\sigma}{\sqrt{n}} + 120 \mid \mu = 120.5\right) = P\left(\frac{\bar{X} - \mu_1}{\sigma/\sqrt{n}} < 1.28 + \frac{120 - 120.5}{\sigma/\sqrt{n}} \mid \mu = 120.5\right) \\ &= \Phi(-1.5933) = 0.0555\end{aligned}$$

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:

$$f_X(x) = \frac{1}{12}x, \quad 1 < x < 5.$$

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

Tempo de secagem (horas)	]1, 3[	[3, 4[	[4, 5[
Frequência absoluta observada (o_i)	36	33	31

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

A: 0.8984 **B:** 0.3978 **C:** 0.6054 **D:** 0.0033

- Como $f_X(x) = \frac{1}{12}x$, para $1 < x < 5$, então a função de distribuição de X é:

$$F_X(x) = \int_1^x \frac{1}{12} u du = \left[\frac{u^2}{24} \right]_1^x = \frac{x^2 - 1}{24}, 1 < x < 5$$

- Teste de ajustamento do Qui-quadrado de Pearson ($n = 100$, $k = 3$):

1. Hipóteses: $H_0 : f_X(x) = \frac{1}{12}x$ versus $H_1 : f_X(x) \neq \frac{1}{12}x$, $1 < x < 5$.

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

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

- $p_1^0 = F_X(3) - F_X(1) = \frac{3^2 - 1}{24} = \frac{8}{24} \approx 0.3333$
- $p_2^0 = F_X(4) - F_X(3) = \frac{4^2 - 1}{24} - \frac{8}{24} = \frac{7}{24} \approx 0.2917$
- $p_3^0 = F_X(5) - F_X(4) = 1 - \frac{15 - 1}{24} = \frac{9}{24} = 0.3750$

i	classe(x) $_i$	o_i	$E_i = 100 \times p_i^0$	$\frac{(o_i - E_i)^2}{E_i}$
1	]1, 3[	36	33.3333	0.2133
2	[3, 4[	33	29.1667	0.5038
3	[4, 5[	31	37.5000	1.1267
total		100	100	$t \approx 1.8438$

3. Valor-p: $P = P(T \geq 1.8438 | H_0) \approx 1 - F_{\chi_{(2)}^2}(1.8438) = 1 - 0.7022 = 0.3978$. Note-se que $P \in (0.3, 0.4) = (1 - F_{\chi_{(2)}^2}(2.408), 1 - F_{\chi_{(2)}^2}(1.833))$.

4. Conclusão: Como valor-p = 0.3978, não se rejeita H_0 para os níveis de significância usuais ($\alpha = 1\%, 5\%$ ou 10%).