

Duração: 120 minutos

Exame Época Recurso - (a)

Justifique convenientemente todas as respostas

Pergunta 1

2 valores

Um sistema eletrónico com duas componentes, 1 e 2, funciona sempre quando ambas as componentes estão operacionais, com probabilidade 0.9 quando apenas a componente 1 está operacional, com probabilidade 0.8 quando apenas a componente 2 está operacional, e nunca quando ambas as componentes estão inoperacionais.

Sabe-se ainda que cada componente está operacional com probabilidade 0.9, independentemente uma da outra. Calcule a probabilidade de o sistema estar em funcionamento.

• **Acontecimentos e probabilidades**

Acontecimento	Probabilidade
$A = \text{"a componente 1 está operacional"}$	$P(A) = 0.9$
$B = \text{"a componente 2 está operacional"}$	$P(B) = 0.9$
$S = \text{"o sistema está em funcionamento"}$	$P(S) = ?$
	$P(S A \cap B) = 1$
	$P(S A \cap \bar{B}) = 0.9$
	$P(S \bar{A} \cap B) = 0.8$
	$P(S \bar{A} \cap \bar{B}) = 0$

• **Cálculo da probabilidade pedida**

$$\begin{aligned}
 P(S) &= P(A \cap B) \times P(S | A \cap B) + P(A \cap \bar{B}) \times P(S | A \cap \bar{B}) \\
 &\quad + P(\bar{A} \cap B) \times P(S | \bar{A} \cap B) + P(\bar{A} \cap \bar{B}) \times P(S | \bar{A} \cap \bar{B}) \quad (\text{lei da probabilidade total}) \\
 &= P(A) \times P(B) \times P(S | A \cap B) + P(A) \times [1 - P(B)] \times P(S | A \cap \bar{B}) \\
 &\quad + [1 - P(A)] \times P(B) \times P(S | \bar{A} \cap B) + [1 - P(A)] \times [1 - P(B)] \times P(S | \bar{A} \cap \bar{B}) \quad (\text{independência de } A \text{ e } B) \\
 &= 0.9 \times 0.9 \times 1 + 0.9 \times (1 - 0.9) \times 0.9 + (1 - 0.9) \times 0.9 \times 0.8 + (1 - 0.9) \times (1 - 0.9) \times 0 \\
 &= 0.963.
 \end{aligned}$$

Pergunta 2

2 valores

Numa determinada população, após contacto com um vírus, cada indivíduo fica infetado com probabilidade 0.8. Suponha que as infeções ocorrem de forma independente.

(a) Qual é a probabilidade de se observarem mais de 14 infeções num total de 20 contactos?

- **V.a. de interesse**

X = Número de infecções num total de 20 contactos.

$$X \sim \text{binomial}(n = 20, p = 0.8).$$

- **Probabilidade pedida**

$$\begin{aligned} P(X > 14) &= 1 - F_X(14) \\ &= P(\text{binom}(20, 0.2) \leq 5) \\ &= 0.8042. \end{aligned}$$

(b) Qual o número esperado de contactos até ocorrer uma infeção?

- **V.a. de interesse**

X = Número de contactos até à primeira infeção.

$$X \sim \text{geométrica}(p = 0.8).$$

- **Cálculo do valor esperado**

$$\begin{aligned} E(X) &= \frac{1}{p} \\ &= \frac{1}{0.8} \\ &= 1.25. \end{aligned}$$

Pergunta 3

2 valores

Um fabricante produz placas de circuito impresso com valor esperado de impedância de 100 ohms. Testes de controle de qualidade indicam que 80% das placas produzidas têm uma impedância entre 95 e 105 ohms. Se a faixa de impedância aceitável é de 90 ohms a 110 ohms, qual é a probabilidade de uma placa ser rejeitada? Suponha uma distribuição normal para as impedâncias.

- **V.a. de interesse**

X = variável aleatória que indica o valor (em ohms) da impedância desse tipo de placa.

- **Distribuição**

$$X \sim \text{normal}(\mu = 100, \sigma^2)$$

- **Sabe-se que**

$$P(95 \leq X \leq 105) = 0.8.$$

- **Queremos determinar**

$$\begin{aligned} P(\text{"placa ser rejeitada"}) &= 1 - P(90 \leq X \leq 110) \\ &= 1 - P\left(\frac{90 - 100}{\sigma} \leq \frac{X - 100}{\sigma} \leq \frac{110 - 100}{\sigma}\right) \\ &= 1 - P(-10/\sigma \leq Z \leq 10/\sigma) \\ &= 1 - \{\Phi(10/\sigma) - \Phi(-10/\sigma)\} \\ &= 2 - 2\Phi(10/\sigma) \\ &= 2 - 2\Phi\left(\frac{10}{3.9015}\right) \\ &\approx 0.01037. \end{aligned}$$

O desvio padrão σ é determinado usando:

$$\begin{aligned}
 P(95 \leq X \leq 105) = 0.8 &\iff P\left(\frac{95-100}{\sigma} \leq \frac{X-100}{\sigma} \leq \frac{105-100}{\sigma}\right) = 0.8 \\
 &\iff P\left(\frac{-5}{\sigma} \leq Z \leq \frac{5}{\sigma}\right) = 0.8 \\
 &\iff 2\Phi(5/\sigma) = 1.8 \\
 &\iff \sigma = \frac{5}{\Phi^{-1}(0.9)} \\
 &\approx 3.9015.
 \end{aligned}$$

Pergunta 4

2 valores

O número de divórcios por indivíduo do sexo masculino adulto, em certa comunidade, foi modelado pela variável aleatória D . As idades desses indivíduos, I , foram classificadas em A="indivíduo com idade ≤ 25 anos"; B="indivíduo com idade entre 26 e 39 anos"; C="indivíduo com idade ≥ 40 anos". A função de probabilidade conjunta do par aleatório (D, I) é dada pela tabela seguinte:

I	D			
	0	1	2	3
A	0.20	a	0	0
B	0.15	0.10	0.05	0
C	0.10	0.05	b	0.05

Sabendo que a probabilidade de indivíduos do sexo masculino com idade entre 26 e 39 anos, de entre os que têm um divórcio, é 40%, prove que $a = 0.10$ e $b = 0.20$. Calcule o número esperado de divórcios de um indivíduo com idade superior ou igual a 40 anos.

• **Cálculo de a e b**

$$\begin{aligned}
 0.40 &= P(I = B | D = 1) \\
 &= \frac{P(I = B, D = 1)}{P(D = 1)} \\
 &= \frac{P(I = B, D = 1)}{a + 0.10 + 0.05} \\
 \Leftrightarrow a + 0.15 &= \frac{0.10}{0.40} \\
 \Leftrightarrow a &= 0.10.
 \end{aligned}$$

$$\sum_{i \in \{A, B, C\}} \sum_{d \in \{0, 1, 2, 3\}} P(I = i, D = d) = 1 \Leftrightarrow b = 0.20.$$

• **Probabilidade $P(I = C) = 0.10 + 0.05 + 0.20 + 0.05 = 0.4$.**

- Probabilidade de D condicionada em $I = C$:

$$P(D = d|I = C) = \frac{P(D = d, I = C)}{P(I = C)} = \begin{cases} \frac{0.10}{0.40} = 0.25, & d = 0, \\ \frac{0.05}{0.40} = 0.125, & d = 1, 3, \\ \frac{0.2}{0.40} = 0.50, & d = 2, \\ 0, & \text{outros casos.} \end{cases}$$

- Valor esperado condicionado

$$E[D|I = C] = \sum_{d=0}^3 d \times P(D = d|I = C) = 1.5.$$

Pergunta 5

2 valores

Suponha que X e Y são duas variáveis aleatórias independentes tais que X tem distribuição normal com valor esperado zero e variância unitária, enquanto que Y tem distribuição normal com valor esperado zero e variância igual a 9. Calcule $P(|X - Y| > 2.5)$.

- Variáveis aleatórias de interesse

$$X \sim \text{normal}(\mu = 0, \sigma^2 = 1);$$

$$Y \sim \text{normal}(\mu = 0, \sigma^2 = 9).$$

X e Y variáveis aleatórias independentes.

- Probabilidade Pedida

$$\begin{aligned} P(|X - Y| > 2.5) &= 1 - P(|X - Y| \leq 2.5) \\ &= 1 - P(-2.5 \leq X - Y \leq 2.5). \end{aligned}$$

- Distribuição de $X - Y$

$$E(X - Y) = E(X) - E(Y) = 0 - 0 = 0$$

$$\text{Var}(X - Y) = \text{Var}(X) + \text{Var}(Y) = 1 + 9 = 10.$$

Logo $X - Y \sim \text{normal}(0, 10)$

- Cálculo da probabilidade pedida

$$\begin{aligned} P(|X - Y| > 2.5) &= 1 - P(-2.5 \leq X - Y \leq 2.5) \\ &= 1 - P\left(\frac{-2.5}{\sqrt{10}} \leq \frac{X - Y}{\sqrt{10}} \leq \frac{2.5}{\sqrt{10}}\right) \\ &\approx 1 - (\Phi(0.79) - \Phi(-0.79)) \\ &= 1 - (\Phi(0.79) - 1 + \Phi(0.79)) \\ &= 2(1 - \Phi(0.79)) \\ &\approx 2(1 - 0.7852) = 0.4296. \end{aligned}$$

Admita que X é uma variável aleatória com função densidade de probabilidade

$$f_X(x; \theta) = \begin{cases} (2x - 1)\theta + 1, & 0 \leq x \leq 1, \\ 0, & \text{outros casos,} \end{cases}$$

onde $\theta \in \{0.9, 1.2, 1.4\}$. A concretização de uma amostra aleatória de dimensão 2 proveniente de X conduziu a $x_1 = 0.3$ e $x_2 = 0.8$. Obtenha a estimativa de máxima verosimilhança de $P(X < 0.3)$.

- Parâmetro desconhecido e espaço paramétrico: θ e $\Theta = \{0.9, 1.2, 1.4\}$
- Seja $\underline{x} = (x_1, x_2)$ uma amostra aleatória de dimensão 2 proveniente da população X .
- **Obtenção da estimativa da máxima verosimilhança de θ**
Passo 1 - Função de verosimilhança

$$\begin{aligned} L(\theta|\underline{x}) &= f_{\underline{X}}(\underline{x}) \\ &\stackrel{X_i \text{ indep}}{=} \prod_{i=1}^2 f_{X_i}(x_i) \\ &\stackrel{X_i \sim X}{=} \prod_{i=1}^2 f_X(x_i) \\ &= \prod_{i=1}^2 [(2x_i - 1)\theta + 1] \\ &= \theta^2(2x_1 - 1)(2x_2 - 1) + \theta[(2x_1 - 1) + (2x_2 - 1)] + 1 \\ &= -0.24 \times \theta^2 + 0.20 \times \theta + 1. \end{aligned}$$

Passo 2 - Maximização e concretização

Como Θ é um conjunto finito, a estimativa de MV de θ obtém-se calculando os vários valores de $L(\theta|\underline{x})$, para $\theta \in \Theta$, e identificando o ponto de máximo - i.e., faz-se por pesquisa ponto a ponto.

$$L(\theta = 0.9|\underline{x}) = 0.9856, \quad L(\theta = 1.2|\underline{x}) = 0.8944, \quad L(\theta = 1.4|\underline{x}) = 0.8096.$$

A função de verosimilhança atinge o máximo quando $\theta = 0.9$, pelo que $\hat{\theta} = 0.9$.

- **Obtenção da estimativa da máxima verosimilhança de $P(X < 0.3)$**

$$P(X < 0.3) = \int_0^{0.3} [(2x - 1)\theta + 1] dx = -0.21 \times \theta + 0.3.$$

Pela propriedade de invariância dos estimadores de máxima verosimilhança, concluímos que a estimativa de MV para $P(X < 0.3)$ é

$$\begin{aligned} \widehat{P(X < 0.3)} &= -0.21 * \hat{\theta} + 0.3 \\ &= -0.21 * 0.9 + 0.3 \\ &= 0.111. \end{aligned}$$

Numa investigação sobre um novo processo de refinação de um certo minério está-se a analisar o teor de lítio no produto refinado. Numa amostra de tamanho 12 observou-se $\sum_{i=1}^{12} x_i = 1140$ e $\sum_{i=1}^{12} x_i^2 = 108900$. Admitindo que o teor de lítio após refinação tem distribuição normal, determine um intervalo de confiança para o teor esperado de lítio a um nível de confiança de 95%.

- **V.a. de interesse**

X = teor de lítio no produto refinado, em percentagem.

- **Situação** $X \sim \text{normal}(\mu, \sigma^2)$.

- **Seleção da variável aleatória fulcral para μ**

$$Z = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{12}}} \sim t_{(11)}.$$

- **Obtenção dos quantis de probabilidade**

$$a_\alpha = \Phi^{-1}(0.025) = -2.201;$$

$$b_\alpha = \Phi^{-1}(0.975) = 2.201.$$

- **Inversão da desigualdade** $a_\alpha \leq Z \leq b_\alpha$

$$P(a_\alpha \leq Z \leq b_\alpha) = 1 - \alpha$$

\Leftrightarrow

$$P\left(-2.201 \leq \sqrt{12} \frac{\bar{X} - \mu}{S} \leq 2.201\right) = 0.95.$$

\Leftrightarrow

$$P\left(\bar{X} - 2.201 \frac{S}{\sqrt{12}} \leq \mu \leq \bar{X} + 2.201 \frac{S}{\sqrt{12}}\right) = 0.95$$

- **Concretização**

$$\begin{aligned} IC_{95\%}(\mu) &= \left[\bar{x} - 2.201 \frac{s}{\sqrt{12}}, \bar{x} + 2.201 \frac{s}{\sqrt{12}} \right] \\ &= \left[95 - 2.201 \frac{7.3855}{\sqrt{12}}, 95 + 2.201 \frac{7.3855}{\sqrt{12}} \right] \\ &= [90.30, 99.69]. \end{aligned}$$

O gerente de uma empresa que fabrica televisores afirma que a proporção de televisores defeituosos produzidos pela sua empresa é igual a 0.045. Para testar esta afirmação, é recolhida uma amostra de televisores, registando-se 60 televisores defeituosos em 1000 inspecionados. Teste a afirmação do gerente, decidindo com base no valor- p .

- **V.a. de interesse**

$X =$ indicador relativo à qualidade da TV = $\begin{cases} 1, & \text{defeituosa,} \\ 0, & \text{não defeituosa,} \end{cases}$

- **Situação**

$X \sim \text{bernoulli}(p)$.

- **Hipóteses**

$$H_0 : p = p_0 = 0.045$$

$$H_1 : p = p_0 \neq 0.045$$

- **Estatística de teste**

$$T = \frac{\bar{X} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \underset{H_0}{\sim} \text{normal}(0, 1)$$

- **Região de rejeição de H_0 (para valores da estatística de teste)**

O teste é bilateral, logo a região de rejeição de H_0 é do tipo $W = (-\infty, -c) \cup (c, \infty)$.

- **Decisão (com base no valor-p)**

Atendendo a que o valor observado da estatística de teste e do valor-p são iguais a:

$$t = \frac{0.06 - 0.045}{\sqrt{0.045 \times (1 - 0.045)/1000}} = 2.28814$$

$$\text{valor-p} \approx 2 \times (1 - \Phi(2.28814)) = 2 \times (1 - 0.9890) = 0.022$$

pelo que H_0 é rejeitada para níveis de significância superiores ou iguais a 0.022.

Pergunta 9	2 valores
-------------------	-----------

Conjetura-se que o número de erros tipográficos em certas páginas web segue uma função de probabilidade Poisson com valor médio igual a 4. Avalie a plausibilidade desta afirmação através do teste de ajustamento do Qui-quadrado (a 1% de significância) explicitando as hipóteses (H_0 e H_1), com base na amostra casual de 50 destas páginas com as seguintes frequências observadas:

nº erros tipográficos por página	≤ 2	3	4	≥ 5
Frequência absoluta observada	12	12	10	16
Frequência absoluta esperada sob H_0	E_1	9.77	9.77	E_4

- **V.a. de interesse**

$X =$ número de erros tipográficos numa página web.

- **Hipóteses**

$H_0 : X \sim \text{poisson}(4)$

$H_1 : X \not\sim \text{poisson}(4)$

- **Estatística de teste**

$$T = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \underset{H_0}{\sim} \chi^2_{(k-1)},$$

onde:

$k =$ No. de classes = 4

$O_i =$ Frequência absoluta observável da classe i

$E_i =$ Frequência absoluta esperada, sob H_0 , da classe i

- **Região de rejeição de H_0** (para valores de T)

Ao efectuarmos um teste de ajustamento do qui-quadrado a região de rejeição de H_0 é um intervalo à direita $W = (c, +\infty)$, onde

$$c = F_{\chi^2_{(k-1)}}^{-1}(1 - \alpha_0) = F_{\chi^2_{(4-1)}}^{-1}(1 - 0.01) = F_{\chi^2_{(3)}}^{-1}(0.99) \stackrel{\text{tabela/calc.}}{=} 11.34.$$

- **Cálculo das frequências absolutas esperadas sob H_0**

$$\begin{aligned} E_i &= n \times p_i^0 \\ &= n \times P(X \in \text{Classe } i \mid H_0) \\ &= n \times P[X \in \text{Classe } i \mid X \sim \text{poisson}(4)] \\ &= n \times \begin{cases} P[X \leq 2 \mid X \sim \text{poisson}(4)], & i = 1, \\ P[X = 3 \mid X \sim \text{poisson}(4)], & i = 2, \\ P[X = 4 \mid X \sim \text{poisson}(4)], & i = 3, \\ P[X \geq 5 \mid X \sim \text{poisson}(4)], & i = 4. \end{cases} \\ &= \begin{cases} 11.90, & i = 1, \\ 9.77, & i = 2, \\ 9.77, & i = 3, \\ 18.56, & i = 4. \end{cases} \end{aligned}$$

• **Decisão**

Assim, temos

$$\begin{aligned}t_0 &= \sum_{i=1}^4 \frac{(O_i - E_i)^2}{E_i} \\&= \frac{(12 - 11.90)^2}{11.90} + \frac{(12 - 9.77)^2}{9.77} + \frac{(10 - 9.77)^2}{9.77} + \frac{(16 - 18.55)^2}{18.56} \\&= 0.8684.\end{aligned}$$

Uma vez que a $t_0 = 0.8655 \notin W_{1\%} = (11.34, +\infty)$ não devemos rejeitar H_0 ao n.s. $\alpha_0 = 1\%$

Pergunta 10

2 valores

Num estudo sobre o efeito do monóxido de carbono (CO) na conservação de carne bovina foram empregues 10 embalagens com diferentes concentrações de CO, x em partes por milhão (ppm), e registados os tempos de conservação da carne, Y em semanas, tendo-se obtido as somas seguintes:

$$\sum_{i=1}^{10} x_i = 7.5, \quad \sum_{i=1}^{10} x_i^2 = 8.81, \quad \sum_{i=1}^{10} y_i = 98.4, \quad \sum_{i=1}^{10} y_i^2 = 1110.48, \quad \sum_{i=1}^{10} x_i y_i = 66.00,$$

onde $[\min_{i=1, \dots, 10} x_i, \max_{i=1, \dots, 10} x_i] = [0.1, 1.9]$. Admita que as variáveis x e Y estão relacionadas de acordo com o modelo de regressão linear simples, $Y = \beta_0 + \beta_1 x + \varepsilon$. Calcule uma estimativa do tempo esperado de conservação da carne para uma concentração de 1 ppm de CO e explique se esse resultado é ou não confiável baseando-se no coeficiente de determinação.

Estimativas de MQ dos parâmetros desconhecidos β_0 e β_1

$$\begin{aligned}\hat{\beta}_1 &= \frac{\sum_{i=1}^{10} x_i y_i - 10 \times \bar{x} \bar{y}}{\sum_{i=1}^{10} x_i^2 - 10 \times \bar{x}^2} \\&= \frac{-7.8}{3.185} \\&= -2.4490.\end{aligned}$$

$$\begin{aligned}\hat{\beta}_0 &= \bar{y} - \hat{\beta}_1 \times \bar{x} \\&= 9.84 - (-2.4489) \times 0.75 \\&= 11.6768.\end{aligned}$$

$$\begin{aligned}E[\widehat{Y} | x = 1] &= \hat{\beta}_0 + \hat{\beta}_1 \\&= 11.6766 - 2.4489 \\&= 9.2278 \text{ semanas.}\end{aligned}$$

Coefficiente de determinação

$$\begin{aligned}r^2 &= \frac{(\sum_{i=1}^{10} x_i y_i - 10 \times \bar{x} \bar{y})^2}{(\sum_{i=1}^{10} x_i^2 - 10 \times \bar{x}^2)(\sum_{i=1}^{10} y_i^2 - 10 \times \bar{y}^2)} \\ &= \frac{(-7.8)^2}{3.185 \times 142.2240} \\ &= 0.1343.\end{aligned}$$

Interpretação do coeficiente de determinação

Apenas cerca de 13% da variação total da variável resposta Y é explicada pela variável x , através do modelo de regressão linear simples ajustado, o que permite concluir que esse modelo não parece ser adequado e, conseqüentemente, a estimativa calculada não é confiável.