

Duração: 120 minutos

Exame de Época de Recurso - Versão A

- 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

Um software especializado numa área de conhecimento recebe documentos para tradução escritos em Inglês, Francês ou Português. A probabilidade de um documento a traduzir conter Inglês, Francês ou Português é de 0.59, 0.49 e 0.26, respetivamente. Alguns destes documentos contêm texto em duas linguas: em Inglês e Francês com probabilidade 0.2, em Português e Inglês com probabilidade 0.1 e, em Português e Francês com probabilidade 0.05. Documentos com as três linguas surgem com probabilidade 0.01. Qual a probabilidade de um documento conter texto em Português, sabendo que contém texto em Francês ou em Inglês?

A: 0.7686 B: 0.8609 C: 0.1591 D: 0.1351

Acontecimentos e probabilidades:

I = documento conter Inglês, $P(I) = 0.59$

F = documento conter Francês, $P(F) = 0.49$

Pt = documento conter Português, $P(Pt) = 0.26$

$P(I \cap F) = 0.2$, $P(I \cap Pt) = 0.1$, $P(Pt \cap F) = 0.05$, $P(I \cap F \cap Pt) = 0.01$

$$P(Pt|F \cup I) = \frac{P(Pt \cap (F \cup I))}{P(F \cup I)} = \frac{P((Pt \cap F) \cup (Pt \cap I))}{P(F \cup I)} = \frac{0.05 + 0.1 - 0.01}{0.49 + 0.59 - 0.2} = 0.1591$$

Pergunta 2

3 valores

A resistência à tração (X , em GPa) de uma fibra de carbono de alto desempenho é modelada como uma variável aleatória contínua com a seguinte função de densidade de probabilidade:

$$f_X(x) = \begin{cases} \frac{3}{8}(x-2)^2, & \text{se } 2 \leq x \leq 4 \\ 0, & \text{caso contrário} \end{cases}$$

(a) Determine a função de distribuição de X .

(1.5)

$$F_X(x) = \int_{-\infty}^x f_X(t) dt = \begin{cases} 0, & \text{se } x \leq 2 \\ \frac{(x-2)^3}{8}, & \text{se } 2 < x \leq 4 \\ 1, & \text{se } x > 4 \end{cases}$$

(b) Calcule a probabilidade de X ser superior a 2.6 sabendo que o valor de X é inferior à mediana.

(1.5)

$$\text{Seja } q_{0.5} \text{ a mediana de } X. P(X > 2.6 | X < q_{0.5}) = \frac{P(2.6 < X < q_{0.5})}{P(X < q_{0.5})} = \frac{F_X(q_{0.5}) - F_X(2.6)}{F_X(q_{0.5})} = \frac{0.5 - 0.027}{0.5} = 0.946$$

Pergunta 3

2 valores

Numa linha de produção automatizada, uma peça que chega ao posto de inspeção pode encontrar-se, com igual probabilidade, em qualquer uma de 12 posições ao longo de uma correia transportadora,

numeradas de 1 a 12. Para recolher a peça, o braço robótico tem de se deslocar até à posição X , em que $X \sim \text{Uniforme}\{1, 2, \dots, 12\}$. Admita que o tempo de reposicionamento do braço, em segundos, é dado por $T(X) = 0.2 + 0.05 \times |X - 6|$. Determine o tempo médio de reposicionamento do braço.

A: 0.35 B: 0.50 C: 0.73 D: 0.15

V.a.:

X = posição da peça ao longo de uma correia transportadora.

$T(X)$ = tempo de reposicionamento do braço, em segundos.

Distribuição de X : Sendo X uma variável aleatória uniforme discreta em $\{1, 2, \dots, 12\}$, tem-se

$$P(X = x) = \frac{1}{12}, \quad x = 1, 2, \dots, 12.$$

Pretende-se calcular o tempo médio de reposicionamento do braço, isto é,

$$E[T(X)] = \sum_{x=1}^{12} T(x) P(X = x) = \sum_{x=1}^{12} (0.2 + 0.05|x - 6|) \frac{1}{12} = 0.2 + \frac{0.05}{12} \sum_{x=1}^{12} |x - 6| = 0.2 + \frac{0.05}{12} \times 36 = 0.35.$$

Isto verifica-se, uma vez que $\sum_{x=1}^{12} |x - 6| = 36$.

Portanto, o tempo médio de reposicionamento do braço é 0.35 segundos.

Pergunta 4

3 valores

Um engenheiro de SRE (*Site Reliability Engineering*) está a monitorizar um contentor que processa tarefas em lotes de, no máximo, 2 pedidos. Seja R o número de pedidos recebidos num intervalo de 10 ms e T o número de pedidos que resultaram em erro de *timeout*. A função de probabilidade conjunta de (R, T) está resumida na seguinte tabela:

$R \setminus T$	0	1	2
0	0.40	0.00	0.00
1	0.25	0.10	0.00
2	0.10	0.10	0.05

(a) Indique o valor do coeficiente de correlação entre R e T .

(2.0)

$$E[R^k] = \sum_{r=0}^2 \sum_{t=0}^2 r^k f_{R,T}(r, t) = \begin{cases} 0.85, & k = 1 \\ 1.35, & k = 2 \end{cases}$$

$$\text{Var}[R] = E[R^2] - (E[R])^2 = 0.6275$$

$$E[T^k] = \sum_{r=0}^2 \sum_{t=0}^2 t^k f_{R,T}(r, t) = \begin{cases} 0.3, & k = 1 \\ 0.4, & k = 2 \end{cases}$$

$$\text{Var}[T] = 0.31$$

$$E[RT] = \sum_{r=0}^2 \sum_{t=0}^2 r t f_{R,T}(r, t) = 0.5$$

$$\text{Cov}[R, T] = E[RT] - E[R]E[T] = 0.245$$

$$\text{Corr}[R, T] = \frac{\text{Cov}[R, T]}{\sqrt{\text{Var}[R]\text{Var}[T]}} = 0.5555$$

(b) Serão R e T variáveis aleatórias independentes?

(1.0)

Como $\text{Corr}[R, T] \neq 0$ ou existe pelo menos um (r, t) tal que $f_{R,T}(r, t) \neq f_R(r)f_T(t)$, e.g., $f_{R,T}(0, 2) = 0 \neq 0.4 \times 0.05 = f_R(0) \times f_T(2)$, então R e T não são variáveis aleatórias independente.

Pergunta 5

2 valores

Um engenheiro de materiais está a medir a porosidade média (θ) de um novo compósito de fibra de carbono. Sabe-se que as medições seguem uma distribuição com média θ e variância unitária. O

engenheiro considera a observação de uma amostra aleatória com n medições e são propostos dois estimadores para a porosidade média: $\bar{X}$, a média amostral, e $T = \frac{n}{n+1} \bar{X}$.

Indique o valor do quociente entre os erros quadráticos médios de $\bar{X}$ e de T para $n = 15$ e $\theta = 3$:

A: 0.61 B: 0.48 C: 0.71 D: 1.04

$$\begin{aligned} E[\bar{X}] &= \theta \text{ e } Var[\bar{X}] = EQM[\bar{X}] = \frac{1}{n} \\ E[T] &= \frac{n}{n+1} \theta \text{ e } Var[T] = \frac{n}{(n+1)^2} \implies EQM[T] = \frac{n+\theta^2}{(n+1)^2} \\ e(\bar{X}, T) &= \frac{EQM[\bar{X}]}{EQM[T]} = \frac{(n+1)^2}{n} \frac{1}{n+\theta^2} = 0.7111111 \end{aligned}$$

Pergunta 6

2 valores

Uma equipa de engenharia do ambiente está a avaliar a concentração de mercúrio nos efluentes de uma unidade industrial antes da sua descarga num rio local. Sabe-se, pelos registos históricos, que a concentração de mercúrio, em $\mu\text{g/L}$, naquele tipo de efluentes segue uma distribuição normal. Para verificar a conformidade com as normas ambientais, foram efetuadas 10 medições independentes num período de 24 horas, tendo-se obtido os seguintes valores:

$$\sum_{i=1}^{10} x_i = 124 \quad \text{e} \quad \sum_{i=1}^{10} x_i^2 = 1540.12$$

Calcule um intervalo de confiança a 95% para a concentração média de mercúrio nos efluentes desta fábrica.

A: (13.12, 13.87) B: (12.02, 12.78) C: (11.12, 11.58) D: (10.32, 10.58)

- Variável aleatória de interesse e distribuição:

X = concentração de mercúrio nos efluentes da fábrica, $X \sim N(\mu, \sigma^2)$

- Método da variável fulcral (estimação intervalar de μ):

- Variável fulcral:

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{(n-1)} = t_{(9)}$$

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

$$P(a < T < b) = 1 - \alpha = 0.95 \text{ e } a = -b \implies b = F_{t_{(9)}}^{-1}(0.975) = 2.262$$

- Desigualdade equivalente:

$$P\left(-b < \frac{\bar{X} - \mu}{S/\sqrt{n}} < b\right) = 0.95 \iff P\left(\bar{X} - b \frac{S}{\sqrt{n}} < \mu < \bar{X} + b \frac{S}{\sqrt{n}}\right) = 0.95$$

- Intervalo aleatório de confiança (IAC) e sua concretização (IC) a 95% para μ :

$$IAC(\mu; 0.95) : \bar{X} \pm b \frac{S}{\sqrt{n}}$$

$$IC(\mu; 0.95) : 12.4 \pm 2.262 \times \frac{0.52915}{\sqrt{10}} \approx (12.02, 12.78),$$

tendo em conta os cálculos auxiliares abaixo.

$$\bar{x} = \frac{\sum_{i=1}^{10} x_i}{n} = \frac{124}{10} = 12.4$$

$$s = \sqrt{\frac{\sum_{i=1}^{10} x_i^2 - n\bar{x}^2}{n-1}} = \sqrt{\frac{1540.12 - 10 \times 12.4^2}{9}} = 0.52915$$

Pergunta 7

3 valores

Seja p a probabilidade de um estudante retirado ao acaso de uma população obter aprovação num dado teste. Considere X a variável indicadora da aprovação de um estudante e $\bar{X}$ a média amostral de uma amostra aleatória de dimensão 100 desta variável.

- (a) Indique, justificando, a distribuição aproximada da estatística $T = \frac{\bar{X} - p}{\sqrt{p(1-p)/100}}$. (1.0)

Das condições do enunciado decorre que $X \sim \text{Ber}(p)$, de modo que

- $E[\bar{X}] \stackrel{i.d.}{=} E(X) = p$
- $\text{Var}(\bar{X}) = \text{Var}(\sum_{i=1}^n X_i)/n^2 \stackrel{ind.}{=} \sum_{i=1}^n \text{Var}(X_i)/n^2 \stackrel{i.d.}{=} \text{Var}[X]/n = p(1-p)/n$,

e da aplicação do Teorema do Limite Central (grandes amostras, i.e., $n = 100$) que

$$T = \frac{\bar{X} - E[\bar{X}]}{\sqrt{\text{Var}[\bar{X}]}} = \frac{\bar{X} - p}{\sqrt{p(1-p)/100}} \stackrel{a}{=} N(0, 1).$$

- (b) Teste, ao nível de significância de aproximadamente 4%, $H_0 : p = 0.75$ versus $H_1 : p > 0.75$, sabendo que 82 dos alunos da amostra obtiveram aprovação no teste. (2.0)

- Estatística de teste:

$$T = \frac{\bar{X} - 0.75}{\sqrt{0.75(1-0.75)/100}} \stackrel{a}{\underset{H_0}{\approx}} N(0, 1)$$

- Região crítica ao nível de significância de 4%: $T > \Phi^{-1}(0.96) = 1.7507$
- Valor observado da estatística de teste: $t = \frac{82/100 - 0.75}{\sqrt{0.75(1-0.75)/100}} \approx 1.6166$
- Conclusão: Não se rejeita H_0 ao nível de significância de 4%.

Pergunta 8

3 valores

No estudo da eficácia de um novo processo de polimento de lentes de alta precisão, um técnico observou o número de vezes, X , que cada lente tem de passar pela máquina até apresentar um acabamento sem imperfeições. Suspeita-se que a probabilidade de sucesso em cada passagem seja constante e igual a 0.7, e que as passagens sejam independentes entre si.

Para averiguar se estas suspeitas são fundamentadas, analisou-se uma amostra de 150 lentes, cujos resultados se encontram na tabela seguinte:

Número de passagens (X)	1	2	3	≥ 4
Frequência observada (o_i)	112	28	8	2

Avalie se o modelo Geométrico de parâmetro $p = 0.7$ é adequado para descrever o número de passagens necessárias para o polimento das lentes. Decida com base no valor- p aproximado.

- $X \sim \text{Geo}(0.7)$, a sua função de probabilidade é $P(X = x) = 0.7 \times (0.3)^{x-1}$, $x = 1, 2, 3, \dots$
- Teste de ajustamento do Qui-quadrado de Pearson $k = 4$ classes.

1. Hipóteses: $H_0 : X \sim \text{Geo}(0.7)$ versus $H_1 : X \not\sim \text{Geo}(0.7)$.

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

- $p_1^0 = P(X = 1) = 0.7$
- $p_2^0 = P(X = 2) = 0.7 \times 0.3 = 0.21$
- $p_3^0 = P(X = 3) = 0.7 \times 0.3^2 = 0.063$
- $p_4^0 = P(X \geq 4) = 1 - (0.7 + 0.21 + 0.063) = 0.027$

2. Estatística do teste: $T = \sum_{i=1}^4 \frac{(O_i - E_i)^2}{E_i} \underset{H_0}{\sim} \chi_{(k-1)}^2 = \chi_{(3)}^2$.

i	classe $(x)_i$	o_i	$E_i = 150 \times p_i^0$	$\frac{(o_i - E_i)^2}{E_i}$
1	1	112	105.00	0.4667
2	2	28	31.50	0.3889
3	3	8	9.45	0.2225
4	≥ 4	2	4.05	1.0377
total		$n = 150$	150	$t \approx 2.1158$

3. Valor-p: $P = P(T > t | H_0) \approx (1 - F_{\chi_{(3)}^2}(2.1158)) = (1 - 0.4513) = 0.5487$. Note-se que $P \in (0.5, 0.6) = (1 - F_{\chi_{(3)}^2}(2.366), 1 - F_{\chi_{(3)}^2}(1.869))$.

4. Conclusão: Rejeita-se H_0 para nível de significância $\alpha > P = 0.5487$ e não se rejeita H_0 no caso contrário. Assim, com base nesta amostra, o modelo Geométrico de parâmetro $p = 0.7$ parece adequado para descrever o número de passagens para os níveis de significância usuais (1%, 5% e 10%).