Probabilidade e Estatística

Duração: **120** minutos **Exame de Época Normal**

Versão A

- As questões de escolha múltipla têm uma e apenas uma resposta correta. Caso assinale uma resposta errada, serão descontados 0.5 valores.
- Se assinalar mais do que uma resposta, será contado como resposta errada.
- As respostas às escolhas múltiplas são dadas na folha de instruções, previamente distribuída.
- Para as restantes questões, terá de justificar convenientemente as respostas.

Pergunta 1 2 valores

Três companhias aéreas com voos domésticos servem o aeroporto Francisco Sá Carneiro. As companhias A, B e C operam, respectivamente, 50%, 30% e 20% dos voos programados. A probabilidade de que um voo da companhia saia no horário programado é de 80%, 65% e 40%, respectivamente para a companhia A, B e C. A probabilidade (considerando 3 casas decimais) de que um voo que acaba de sair no horário programado seja da companhia C é:

A: 0.080 B: 0.119 C: 0.216 D: 0.400

- **Acontecimentos e probabilidades**

Acontecimento	Probabilidade
A = "voo ser da companhia A"	$P(A) = 0.5$
B = "voo ser da companhia B'	$P(B) = 0.3$
C = "voo ser da companhia C"	$P(C) = 0.2$
D = "voo sair no horário programado	$P(D\|A) = 0.80$, $P(D\|B) = 0.65, P(D\|C) = 0.40$

- **Cálculo da probabilidade pedida** (usando teorema de Bayes e teorema das probabilidades totais)

$$\begin{aligned} P(C \mid D) &= \frac{P(C) \times P(D \mid C)}{P(A) \times P(D \mid A) + P(B) \times P(D \mid B) + P(C) \times P(D \mid C)} \\ &= \frac{0.2 \times 0.40}{0.5 \times 0.80 + 0.3 \times 0.65 + 0.2 \times 0.40} = 0.119 \end{aligned}$$

Pergunta 2 2 valores

Uma empresa leva, em média, 18 meses para executar uma obra. O tempo de execução, em meses, para cada obra desta empresa segue uma distribuição exponencial. Para uma obra que tenha sido adjudicada, que prazo de conclusão deve ser fixado para que a probabilidade do seu cumprimento seja de 0.80?

- **V.a. de interesse**

$$X = \text{tempo de execução de uma obra (em meses)}$$

- **Distribuição**

$$X \sim Exp(\lambda)$$

com $\mathbb{E}[X] = 18$ pelo que $\lambda = \frac{1}{18}$.

- **Prazo de entrega pretendido,** K

$$K = F^{-1}_{Exp(1/18)}(0.80) \Leftrightarrow \int_0^K \frac{1}{18} e^{-\frac{x}{18}} dx = 0.80 \Leftrightarrow 1 - e^{-K/18} = 0.80 \Leftrightarrow K = 28.970$$

Pergunta 3	2 valores

Um investigador precisa de 2 homens daltónicos para uma experiência e, para tal, vai sortear ao acaso candidatos de uma grande lista de voluntários para avaliar a presença do daltonismo. Considerando que 8% dos homens desta lista são daltónicos, calcule a probabilidade de o investigador precisar de fazer 10 sorteios até encontrar 2 homens daltónicos.

- **V.a. de interesse**

$$X = \text{número de observações até encontrar o 2º homem daltónico}$$

- **Probabilidade pedida**

$$P(X = 10) = 9 \times (1 - 0.08)^8 \times 0.08^2 = 0.0296$$

Pergunta 4	2 valores

Uma empresa vende caixas com biscoitos e envia-as pelo correio, pagando um extra nos portes de envio se o peso total da caixa for superior a 2508 g. Cada caixa leva 100 biscoitos e o peso da embalagem é desprezível. Se soubermos que o peso de cada biscoito é variável mas que em média pesa 25g, com um desvio padrão de 8g, determine um valor aproximado para a probabilidade de ter que pagar um extra num envio.

- **V.a. de interesse**

$$X = \text{peso, em g, de um biscoito}$$
$$S = \text{peso da caixa } = \sum_{i=1}^{100} X_i$$

com $\{X_i, i = 1, \ldots, 100\}$ i.i.d. a X.

- **Momentos**

$$E[X] = 25, Var[X] = 8^2$$

pelo que

$$E[S] = 100 \times 25 = 2500, Var[S] = 100 \times 8^2 = 6400$$

- **Probabilidade pedida**: Pelo teorema do Limite Central:

$$P(S > 2508) = 1 - P(S \leq 2508) = 1 - P\left(\frac{S - 2500}{\sqrt{6400}} \leq \frac{2508 - 2500}{\sqrt{6400}}\right)$$
$$= 1 - P(Z \leq 0.1) \approx 1 - \Phi(0.1) = 1 - 0.540 = 0.460$$

Pergunta 5	2 valores

A função de probabilidade conjunta de (X, Y) é dada por:

$$P(X = x, Y = y) = \frac{x + y}{32}, \quad x = 1, 2; y = 1, 2, 3, 4$$

O valor esperado $\mathbb{E}[X|Y \geq 3]$ é igual a:

A: 1.250 B: 1.455 C: 1.752 D: 1.550

- **Probabilidades marginais**

$$P(X = x) = \sum_{y=1}^{4} P(X = x, Y = y) = \sum_{y=1}^{4} \frac{x + y}{32} = \frac{x}{8} + \frac{1 + 2 + 3 + 4}{32}, \quad x = 1, 2$$
$$P(Y = y) = \sum_{x=1}^{2} P(X = x, Y = y) = \frac{y}{16} + \frac{1 + 2}{32}, \quad y = 1, 2, 3, 4.$$

- $P(Y \geq 3) = \frac{3+4}{16} + \frac{3}{32} = \frac{20}{32}$

- **Probabilidade condicionada de X em $Y \geq 3$**

$$P(X = x | Y \geq 3) = \frac{P(X = x, Y \geq 3}{P(Y \geq 3} = \begin{cases} \frac{9}{20} & x = 1 \\ \frac{11}{20} & x = 2 \end{cases}$$

- $\mathbb{E}[X|Y \geq 3] = \sum_{x=1}^{2} x P(X = x | Y \geq 3) = 1 \times \frac{9}{20} + 2 \times \frac{11}{20} = 1.550$

Pergunta 6	2 valores

O número diário de erupções solares de classe M (erupções de média intensidade) pode ser considerado como uma variável aleatória de Poisson, de valor esperado μ. Em 20 dias foram contabilizadas 10 erupções de classe M. A estimativa de máxima verosimilhança da probabilidade de num dia haver uma erupção de classe M, sabendo que nesse dia ocorreram erupções é:

A: 0.500 B: 0.303 C: 0.393 D: 0.771

Seja X_i a v.a. que indica o número de erupções solares no dia i, com $X_i \sim \text{Poisson}(\mu)$.

- **função de verosimilhança de** μ, com base numa amostra de dimensão 20

$$L(\mu; x_1, \ldots, x_{20}) = \prod_{i=1}^{20} \frac{e^{-\mu}\mu^{x_i}}{x_i!} = \frac{e^{-20\mu}\mu^{\sum_{i=1}^{20} x_i}}{\prod_{i=1}^{20} x_i!}$$

- **Log Verosimilhança e derivadas**

$$\ln L(\mu; x_1, \ldots, x_{20}) = -20\mu + \sum_{i=1}^{20} x_i \ln\mu - \ln \prod_{i=1}^{20} x_i!$$

$$\left(\ln L(\mu; x_1, \ldots, x_{20})\right)' = -20 + \frac{\sum_{i=1}^{20} x_i}{\mu}$$

$$\left(\ln L(\mu; x_1, \ldots, x_{20})\right)'' = -\frac{\sum_{i=1}^{20} x_i}{\mu^2} < 0, \quad \forall \mu > 0$$

- **Estimativas**

$$\hat{\mu} = \frac{\sum_{i=1}^{20} x_i}{20} = 0.5$$

$$P\,(\widehat{X = 1|X} > 0) = \left(\frac{\widehat{P(X=1)}}{1 - P(X=0)}\right) = \frac{0.5e^{-0.5}}{1 - e^{-0.5}} = 0.771$$

pelo princípio de invariância dos estimadores de máxima verosimilhança.

Pergunta 7	2 valores

A quantidade de açúcar (em grama) na calda de pêssegos em lata tem distribuição normal. É extraída uma amostra de $n = 10$ latas que resulta num desvio padrão amostral $s = 4.8$. Determine o intervalo de confiança a 95% para a variância populacional, σ^2.

- **V.a. de interesse**

$$X \sim N(\mu, \sigma^2)$$

com μ e σ desconhecidos.

- **Variável fulcral**

$$T = \frac{9S^2}{\sigma^2} \sim \chi^2_{(9)}$$

- **Quantis**

$$F^{-1}_{\chi^2_{(9)}}(0.025) = 2.70; \quad F^{-1}_{\chi^2_{(9)}}(0.975) = 19.02;$$

- **ICA**$_{0.95}(\sigma^2)$

$$\left[\frac{9S^2}{19.02}, \frac{9S^2}{2.70}\right]$$

- **IC**$_{0.95}(\sigma^2)$

$$\left[\frac{9 \times 4.8^2}{19.02}, \frac{9 \times 4.8^2}{2.70}\right] = [10.90, 76.8]$$

Pergunta 8	2 valores

Seja X uma variável aleatória com distribuição normal, de variância 25. Foi recolhida uma amostra de 25 observações, que conduziu a uma média amostral igual a 4. O valor-p para o teste $H_0 : \mu = 4.5$ versus $H_1 : \mu > 4.5$ é:

A: 0.3108 B: 0.6915 C: 0.3085 D: 0.6171

- **Variável Fulcral**

$$T = \bar{X} - \mu \sim N(0,1)$$

- **Estatística de Teste**

$$T_0 = \bar{X} - 4.5 \sim N(0,1)$$

- **Valor observado da estatística de Teste**

$$t_0 = -0.5$$

- **Valor-p**

$$p = 1 - \Phi(-0.5) = 0.6915$$

pelo que a hipótese não é rejeitada para os valores usuais de significância.

Pergunta 9 2 valores

Um modelo de automóvel é vendido em quatro versões: SX, LX, GLX, GTX. Foi feita uma campanha publicitária para melhorar as vendas das versões GLX e GTX, e posteriormente foi verificada a escolha das versões em 500 vendas escolhidas ao acaso. Os resultados foram:

Versão	SX	LX	GLX	GTX
Unidades vendidas	210	125	105	60

De acordo com o fabricante, com esta campanha as vendas deste modelo deve distribuir-se pelas versões do seguinte modo: 40% de SX, 30%de LX, 20% de GLX e 10% de GTX. Teste se o fabricante está correcto, assumindo um nível de significância de 10%.

- **Variável aleatória de interesse**

$$X = \text{(indicatriz) da versão do modelo automóvel} = \begin{cases} 1 & \text{versão SX} \\ 2 & \text{versão LX} \\ 3 & \text{versão GLX} \\ 4 & \text{versão GTX} \end{cases}$$

- **Hipóteses**

$$\mathrm{H}_0 : P(X = i) = \begin{cases} 0.40 & x = 1 \\ 0.30 & x = 2 \\ 0.20 & x = 3 \\ 0.10 & x = 4 \end{cases}$$

$$\mathrm{H}_1 : \text{outra distribuição para } X$$

- **Estatística de teste**

$$T = \sum_{i=1}^{4} \frac{(E_i - O_i)^2}{E_i} \sim_a \chi^2_{(3)}$$

- **Valor observado da estatística de teste**

$$t_0 = \frac{(210-200)^2}{200} + \frac{(125-150)^2}{150} + \frac{(105-100)^2}{100} + \frac{(60-50)^2}{50} = 6.92$$

- **Região de rejeição**

  Rejeitar H_0 para valores observados da estatística de teste maiores que $F^{-1}_{\chi^2_{(3)}}(0.90) = 6.251$

- **Decisão**

  Rejeitar a hipótese para $\alpha = 10\%$.

Pergunta 10	2 valores

O gestor de marketing duma grande cadeia de supermercados quer determinar o efeito do espaço de prateleiras nas vendas de comida para animais. Seleccionou uma amostra de 12 lojas para as quais recolheu dados relativos ao espaço de prateleira (em metro, x_i, com valores no intervalo $(5,20)$) e as vendas numa semana ($\times 10^3$ €, Y_i), tendo obtido os seguintes resultados:

$$\sum_{i=1}^{12} x_i = 150, \quad \sum_{i=1}^{12} x_i^2 = 2250, \quad \sum_{i=1}^{12} y_i = 28.5, \quad \sum_{i=1}^{12} y_i^2 = 70.69, \quad \sum_{i=1}^{12} x_i y_i = 384$$

Admitindo a validade do modelo de regressão linear simples, determine a estimativa do aumento médio das vendas numa semana quando se aumenta o espaço de prateleira em 2 metros.

- **Estimativa dos parâmetro**

$$\widehat{\beta}_1 = \frac{\sum x_i y_i - n\overline{xy}}{\sum x_i^2 - n\overline{x}^2} = \frac{384 - 12 \times (150/12) \times (28.5/12)}{2250 - 12 \times (150/12)^2} = 0.074$$

- **Aumento médio das vendas numa semana quando se aumenta o espaço de prateleira em 2 metros**

$$E[Y|X = x+2] - E[Y|X = x] = 2\beta_1$$

- **Estimativa do aumento médio das vendas numa semana quando se aumenta o espaço de prateleira em 2 metros**

$$2\widehat{\beta_1} = 2 \times 0.074$$

desde que x e $x+2$ estejam na gama de valores observados para a variável x (intervalo (5,20)).