Instituto Superior de Economia e Gestão

Universidade de Lisboa

Licenciaturas em Economia e em Finanças

Estatística II – Teste intercalar - 08/04/2016 – Duração 1 hora

Nome:____ Número:____________

Notas: Certifique-se de que o seu telemóvel está desligado. Se não estiver, é motivo suficiente para anulação da prova. As perguntas de escolha múltipla valem 1 valor; respostas erradas são penalizadas em 0.25. Caso nada seja dito em contrário, utilize um nível de significância de 5% nos testes de hipóteses que efetuar. Fundamente e formalize devidamente todas as respostas. Pode usar a última página para continuar qualquer questão.

Espaço reservado para classificações

1. Considere que o valor diário em u.m. que entra na caixa de um dado supermercado tem distribuição Normal de parâmetros desconhecidos. Numa amostra casual de 20 dias observou-se que em média entrou em caixa 126 u.m. com variância 361.

   a) Sabendo que o estimador da Máxima verosimilhança para μ é $\bar{X}$ e para σ^2 é S^2 estime pela máxima verosimilhança a probabilidade de num dia entrarem mais de 155 u.m. em caixa. Justifique devidamente a sua resposta. [**1.5**]

Seja X o montante diário que entra na caixa do supermercado, em u.m., onde $X \sim N(\mu, \sigma^2)$

Amostra: $n = 20$, $\bar{x} = 126$, $s^2 = 361$.

Estimadores MV: $\hat{\mu} = \bar{X}$, $\hat{\sigma}^2 = S^2$

Pretende-se a estimativa MV para $P(X > 155) = \tau(\mu, \sigma^2)$, onde τ é função biunívoca de μ e σ^2. Então, pela propriedade da invariância dos estimadores MV sabe-se que $\hat{\tau}(\mu, \sigma^2) = \tau(\hat{\mu}, \hat{\sigma}^2)$ e logo conclui-se que:

$$\hat{P}(X > 155) = P\left(\frac{X - \hat{\mu}}{\hat{\sigma}} > \frac{155 - 126}{\sqrt{361}}\right) = P(Z > 1.53) = 1 - \Phi(1.53) = 1 - 0.9370 = 0.063$$

b) O proprietário do supermercado tem como objetivo de negócio que em média entre na caixa diariamente pelo menos 135 euros. Perante a evidência dada na amostra acha que este objetivo é de rejeitar considerando um teste de dimensão 1%? [**1.5**]

Hipóteses:

$$H_0: \mu \geq 135 \quad vs \quad H_1: \mu < 135$$

Estatística de Teste (sob H_0):

$$T = \frac{\bar{X} - \mu_0}{S'/\sqrt{n}} \sim t(n-1) \to t(19)$$

Região de rejeição:

$$W_{0.01} = \{t \ : t \ \ < -2.539\}$$

Valor observado da estatística teste:

$$t_{obs} = \frac{\bar{x} - \mu_0}{s'/\sqrt{n}} = \frac{\bar{x} - \mu_0}{s/\sqrt{n-1}} = \frac{126 - 135}{\sqrt{361}/\sqrt{19}} \approx -2.065$$

Porque $t_{obs} \notin W_{0.01}$, não se rejeita H_0 ao nível de 1%. Perante a evidência dada na amostra, podemos concluir que o objetivo de negócio não deve ser rejeitado.

Alternativamente, pelo valor-p, tem-se:

$p_{obs} = P(T \leq t_{obs}|H_0) = P(T \leq -2.065) \approx 0.0264 > \alpha = 0.01 \Rightarrow$ Não rejeitar H_0 a 1%...

2. Sejam T_1 e T_2 dois estimadores para θ. Sabe-se que T_1 é centrado e $E(T_2) = \theta + k$. Escolha a opção Verdadeira.

- ☐ Se $k = 0$ então necessariamente $T_1 = T_2$.
- ☐ Se $k < 0$ e $Var(T_1) = Var(T_2)$ então T_2 é preferível a T_1 em Erro Quadrático Médio.
- ☒ Se $k > 0$ e $Var(T_2) < Var(T_1) - k^2$ então T_2 é preferível a T_1 em Erro Quadrático Médio.
- ☐ Nenhuma das opções anteriores é verdadeira.

3. Seja uma população com distribuição Normal e o teste de hipóteses, $H_0 : \sigma^2 = \sigma_0^2$ contra $H_1 : \sigma^2 > \sigma_0^2$. Para uma amostra de 25 observações foram definidas as seguintes regiões de rejeição: $W_1 = \{q : q < 12.401\}$ e $W_2 = \{q : q > 39.364\}$ com q a estatística de teste com distribuição $\chi^2(24)$ sob H_0.

- ☐ A probabilidade do erro de 1ª espécie com W_1 é igual à probabilidade do erro de 1ª espécie com W_2 e é igual a 5%.
- ☒ W_2 é preferível a W_1 porque com W_2 a probabilidade do erro de 2ª espécie é menor para o mesmo α.
- ☐ W_2 é preferível a W_1 porque com W_1 a probabilidade do erro de 1ª espécie é maior.
- ☐ Se H_0 é verdadeira então W_1 é preferível a W_2 .

4. Com uma amostra de 100 observações obteve-se $\bar{x} = 25$, $s' = 5$, e o seguinte intervalo de confiança para μ : $(23.837 ; 26.163)$. O nível de confiança associado a este intervalo é aproximadamente igual a,

- [x] 98%
- [ ] 95%
- [ ] 99%
- [ ] 90%

5. O gasto mensal de energia de uma empresa é uma variável aleatória com distribuição Normal com média desconhecida e variância igual a 900. O gestor desta empresa pretende avaliar se estes gastos, em média, não são superiores a 220 através do teste de hipóteses $H_0 : \mu \leq 220$ contra $H_1 : \mu > 220$. O gestor decide que o seu objetivo não é atingido se numa amostra casual, de dimensão 15, se verificar $\bar{x} > 233$. Calcule a probabilidade do gestor não cometer um erro com este teste quando o verdadeiro gasto médio mensal é igual a 240. [**2.0**]

Seja X o gasto mensal de energia, em u.m., onde $X \sim N(\mu, \sigma^2)$ com $\sigma^2 = 900$

Quando o verdadeiro gasto médio mensal é $\mu = 240$, está-se a admitir H_0 falsa. Assim, o gestor não comete erro se efetivamente rejeitar H_0, ou seja, pretende-se calcular a seguinte probabilidade:

$$P(\text{rejeitar } H_0 | H_0 \text{ falsa}) \rightarrow \text{Potência do teste } (\beta)$$

Porque o gestor rejeita o objetivo (H_0) quando $\bar{x} > 233$, tem-se então:

$$\beta = P(\bar{X} > 233 | \mu = 240) = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} > \frac{233 - 240}{\sqrt{900}/\sqrt{15}}\right) \approx P(Z > -0.90) = \Phi(0.90) = 0.8159$$

6. Um estudo de mercado pretende avaliar se existem diferenças entre homens e mulheres sobre a preferência entre ler os livros impressos em papel ou em formato eletrónico pdf. Numa sondagem foram entrevistadas 200 pessoas escolhidas aleatoriamente, tendo sido obtidos os seguintes resultados:

	Prefere ler livro em papel	Prefere ler pdf	Total
Mulher	85	35	120
Homem	55	25	80
Total	140	60	200

Será de concluir que as preferências do formato de leitura dependem do sexo da pessoa? Justifique através de um teste de hipóteses adequado. [**2.0**]

<u>Teste de Independência</u> $H_0: p_{ij} = p_{io}p_{oj} \quad (i,j = 1,2) \quad vs \quad H_1: \exists(i,j): p_{ij} \neq p_{io}p_{oj} \quad (i,j = 1,2)$

Estatística de teste (sob H_0):

$$Q = \sum_{i=1}^{2}\sum_{j=1}^{2}\frac{\left(N_{ij} - fe_{ij}\right)^2}{fe_{ij}} \overset{a}{\sim} \chi^2(1)$$

Região de rejeição ($\alpha = 0.05$):

$$W_{0.05} = \{q_{obs}: q_{obs} > 3.841\}$$

Cálculo das frequências esperadas:

$$fe_{11} = \frac{120 \times 140}{200} = 84\,; \quad fe_{12} = \frac{120 \times 160}{200} = 36$$
$$fe_{21} = \frac{80 \times 140}{200} = 56\,; \quad fe_{22} = \frac{80 \times 60}{200} = 24$$

Valor observado da estatística teste:

$$q_{obs} = \frac{(85-84)^2}{84} + \frac{(35-36)^2}{36} + \frac{(55-56)^2}{56} + \frac{(25-24)^2}{24} \approx 0.0992$$

Porque $\notin W_{0.05}$, não se rejeita H_0 ao nível de 5%. Evidência estatística favorável à independência entre a preferência do formato de leitura e o sexo da pessoa.

Alternativamente, pelo valor-p, tem-se:

$p_{obs} = P(Q \geq q_{obs}|H_0) = P(Q \geq 0.0992) \approx 0.7528 \Rightarrow$ Não rejeitar H_0 aos níveis de significância habituais...