ESTATÍSTICA II - 2º Ano/2º sem. Gestão do Desporto   Época Recurso   03. 07. 17

2horas. (20 valores)

Nome:____

Espaço reservado para classificações

1 a) (10)	2 a) (20)	3 a) (20)	*4 a) (10)*	*4.c)(20)*	
1 b) (20)	2 b) (20)	3 b) (20)	*4 b)* (20)	*4.d)*(20)	*5.* (20)

T:

Atenção: 1. As folhas EXCEL no écran do computador tem os dados para a resolução de todas as questões do exame.

2. Intervalos de confiança, ensaios de hipóteses e regressão tem de ser feitos usando o EXCEL.

3. As questões não podem ser respondidas usando o método de tentativa e erro.

4. Devem apresentar na folha de exame a formalização e Justificação dos cálculos efectuados no EXCEL.

5. Devem fazer os cálculos no ficheiro EXCEL em folhas separadas para cada questão.

1. O departamento de risco de uma instituição bancária considera que o montante de crédito, em dezenas de milhares de euros, concedido por essa instituição é uma variável aleatória com a seguinte função densidade:

$$f_X(x) = \frac{2}{\alpha^2}x \quad 0 < x < \alpha \quad \alpha > 0$$

Em que $E(X) = \frac{2}{3}\alpha;\ \ Var(X) = \frac{\alpha^2}{18}$.

a) Deduza o estimador pelo método dos momentos para α.

+

$E(X) = \bar{X} \Leftrightarrow \frac{2}{3}\tilde{\alpha} = \bar{X} \Leftrightarrow \tilde{\alpha} = \frac{3}{2}\bar{X}$ é o estimador pelo método dos momentos para α

b) Se se considerar um outro estimador $T(X_1, X_2, \cdots, X_n) = \bar{X}$ para α. Estude a eficiência relativa deste estimador em relação ao encontrado na alínea a).

$$Var(\bar{X}) = \frac{\sigma^2}{n} = \frac{\frac{\alpha^2}{18}}{n} = \frac{\alpha^2}{18n};\quad Var(\tilde{\alpha}) = Var\left(\frac{3}{2}\bar{X}\right) = \frac{9}{4} * Var(\bar{X}) = \frac{9}{4} * \frac{\alpha^2}{18n}$$

$$\frac{Var(\bar{X})}{Var(\tilde{\alpha})} = \frac{\frac{\alpha^2}{18n}}{\frac{9}{4}*\frac{\alpha^2}{18n}} = \frac{4}{9} < 1 \Rightarrow Var(\bar{X}) < Var(\tilde{\alpha}) \Rightarrow \bar{X}$$ é um estimador

relativamente mais eficiente que $\tilde{\alpha}$.

2. Uma empresa que fabrica computadores recolheu uma amostra de 300 computadores para a qual registou, entre outras variáveis a velocidade do processador.

a) Determine a estimativa por intervalos para a média da velocidade do processador dos computadores produzidos pela fábrica para um grau de confiança de 90%. Interprete o resultado obtido.

Variável Fulcral - $\frac{\bar{X}-\mu}{S'/\sqrt{n}} \sim t_{(n-1)}$

$$IC_{\mu}^{90\%} = (4.30, 4.59)$$

Se se calcularem um grande número de intervalos de confiança com amostras da mesma dimensão e idêntico grau de confiança, em 90% deles a média pertencerá ao intervalo de confiança.

b) Qual a margem de erro do intervalo de confiança que calculou? Se pretendesse reduzir para metade a margem de erro, mantendo o grau de confiança, qual deveria ser a dimensão da amostra a selecionar?

Margem de erro $\epsilon = 0.1427$ (do output de excel)$= t_{\alpha/2} \frac{s'}{\sqrt{n}}$

$$t_{0.1/2} = t_{0.05}: P\left(T_{(299)} > t_{0.05}\right) = 0.05 \Rightarrow t_{0.05} = 1.65 \ (output\ do\ excel)$$

$$0.1427 = 1.65\frac{1.498}{\sqrt{n}} \Leftrightarrow \sqrt{n} = \frac{1.65 * 1.498}{0.1427} = 17.321 \Leftrightarrow n = 300$$

3. A fábrica em questão garante que a média da velocidade do processador dos computadores por ela fabricados é superior a 5.

a) Com base num teste adequado ao nível de 1 % diga se é de pôr em causa a garantia do fabricante? (Nota: formule as hipóteses justificando a escolha, indique a estatística teste e respectiva distribuição)

$$H_0: \mu \geq 5 \quad contra \quad H_1: \mu < 5$$

Porque a afirmação que se pretende testar deve aparecer sempre na hipótese nula.

Estatística teste: $\frac{\bar{X}-\mu}{S'/\sqrt{n}} \sim t_{(n-1)}$

Valores obtidos no output do EXCEL

$$W = \left\{t: \underbrace{t_{obs} < -2.3389}_{porque\ a\ RR\ está\ sempre\ do\ aldo\ da\ alternativa}\right\} \Rightarrow t_{obs} = -6.416 \in W$$

$$valor - p = P\left(T_{(299)} < -6.416\right) = 0.0000$$

Então rejeita-se H_0 o que significa que se deve pôr em causa a garantia.

b) Determine a potência do ensaio se a verdadeira média da velocidade do processador dos computadores fabricados por esta fábrica for de 4.5.

. $\beta(4.5) = P(Rejeitar\ H_0|H_0 falsa) = P\left(Rejeitar\ H_0|\mu = 4.5\right) = P(\bar{X} < -2.3389)$

$$= P\left(\frac{\bar{X} - \mu}{S'/\sqrt{n}} < 4.2973\right) = P\left(T_{(299)} < 4.2973\right) = 0.9999$$

4. O CEO da empresa pretende conhecer o impacto no ***preço*** dos computadores de variáveis como a capacidade do disco rígido do computador em gigabytes (***discrig.***), memória RAM do computador em gigabytes (***ram***), velocidade do processador do computador em gigahertz (***vel***) e sistema operativo (***Windows*** = 1 se sistema operativo é Windows, 0 no caso contrário). Para tal adoptou o seguinte modelo:

$$\ln(\boldsymbol{preço_i}) = \beta_0 + \beta_1 \boldsymbol{discrig}_i + \beta_2\, \boldsymbol{ram}_i + \beta_3 \ln(\boldsymbol{vel}_i) + \ \beta_4\, \boldsymbol{Windows}_i + u_i$$

a) Estime o modelo.

b) Interprete as estimativas dos coeficientes associados aos regressores $\boldsymbol{lnvel}$ e $\boldsymbol{Windows}$.

b_3- tudo o resto constante, um acréscimo de 1% da velocidade do processador do computador faz o preço crescer, em média, 28.33%

b_4 –um computador com sistema operativo ***Windows*** tem um preço 1.2% superior a um computador sem sistema operativo ***Windows***

c) Realizando o teste adequado diga se está de acordo com a afirmação de que, tudo o resto constante, o acréscimo de um gigabyte de memória RAM aumenta, em média, o preço dos computadores em menos de 1%.

$$H_0: \beta_2 \leq 0.01 \quad contra \quad H_1: \beta_2 > 0.01$$

Estatística teste: $\frac{b_2 - \beta_2}{s_{b_2}} \sim t_{(n-k-1)}$

$$t_{obs} = \frac{0.010511 - 0.01}{0.002428} = 0.2106 \Rightarrow valor - p = P\left(T_{(295)} > 0.2106\right) = 0.42$$

$$\Rightarrow valor - p > 0.05 \Rightarrow não\ se\ rejeita\ H_0$$

d) Estime o preço médio de um computador com 400 gigabytes de capacidade do disco rígido, 16 gigabytes de memória RAM, 8 gigahertz de velocidade do processador e sistema operativo Windows.

$$E(\ln \widehat{\boldsymbol{preço}}) = 6.924993+0.001613*400+0.010511*16+0.283313*\ln(8)+0.012553$$

$$=8.33995$$

$$\boldsymbol{E}(\widehat{\boldsymbol{preço}}) = \boldsymbol{e}^{8.33995} = 4187.89$$

5. Um inquérito a 500 portugueses com mais de 18 anos sobre a compra “online” de produtos de supermercado forneceu os seguintes resultados:

	Compra “online”	Não compra “online”
(18, 30]	73	78
(30, 50]	48	111
>50	20	170

A um nível de significância de 5% existe evidência empírica para afirmar que a compra “online” é independente da idade?

$$H_0: p_{ij} = p_{i.} * p_{.j} \ \forall (i,j) \quad \boldsymbol{contra} \ \exists (i,j): p_{ij} \neq p_{i.} * p_{.j}$$

$$Q_{Obs.} = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{\left(\boldsymbol{n_{ij}} - n.\widehat{p_{\iota}}\,\widehat{p_{J}}\right)^2}{n\,\widehat{p_{\iota}}\,\widehat{p_{J}}} = 59.88473$$

$$Valor - p = P\left(\chi^2_{(2-1)(3-1)} > 59.88473\right) \approx 0$$

Então, rejeita-se H_0, a compra “online” não é independente da idade