Estatística II

Licenciatura em Gestão

1º semestre – 2015/2016
14/01/2016

Nome____ **N**^{o}__________

Espaço reservado a classificações

A utilização do telemóvel, em qualquer circunstância, é motivo suficiente para a anulação da prova.

Perguntas de escolha múltipla: apenas uma opção é correta
cada resposta certa vale **1** valor
cada resposta errada vale **-0.25** valores

Pergunta Verdadeiro/Falso: cada resposta certa vale **0.25** valores
cada resposta errada vale **-0.25** valores

Se necessitar de espaço utilize a última página do enunciado, indicando com clareza a respetiva questão.

É expressamente proíbido destacar as folhas do enunciado.

TESTE I

1. Para cada afirmação assinale se esta é verdadeira **(V)** ou falsa **(F)**. (1.0)

	V	**F**
Num teste estatístico de dimensão $\alpha = 0.05$, rejeita-se H_0 sempre que $t < 0.05$, em que t é o valor observado da estatística de teste.		
Seja $T = \dfrac{X_1 + X_2 + \sum_{i=1}^{n} X_i}{n+2}$ um estimador para o valor esperado de uma variável aleatória X, $\mu = E(X)$, com base numa amostra casual de dimensão n. Então T é estimador centrado de μ.		
Num teste estatístico de dimensão $\alpha = 0.1$ não se rejeita H_0 se o valor-p for 0.09.		
Uma estatística é uma variável aleatória.		

2. Num restaurante foram selecionadas ao acaso 61 faturas com indicação de NIF, tendo-se verificado que a soma dos montantes foi de 1279 euros e a sua variância corrigida $s'^2 = 62$. Assume-se que o montante, em euros, de faturas com NIF tem distribuição normal.

(a) Pretende-se testar, ao nível de significância de 5%, se o montante médio de faturas com NIF é 21 euros. (1.0)
A região crítica do teste estatístico correspondente será:

☐ $W = \left\{(x_1, \ldots, x_n) : \bar{x} < 21 - 2\sqrt{1.0164}\right\}$

☐ $W = \left\{(x_1, \ldots, x_n) : \bar{x} > 21 + 2\sqrt{1.0164}\right\}$

☐ $W = \left\{(x_1, \ldots, x_n) : \bar{x} < 21 - 1.645\sqrt{1.0164} \text{ ou } \bar{x} > 21 + 1.645\sqrt{1.0164}\right\}$

☐ $W = \left\{(x_1, \ldots, x_n) : \bar{x} < 21 - 2\sqrt{1.0164} \text{ ou } \bar{x} > 21 + 2\sqrt{1.0164}\right\}$

(b) Com base na alínea anterior, o que pode concluir sobre a afirmação "o montante médio das faturas com NIF é 21"? Justifique. (1.0)

(c) Na mesma loja foram selecionadas ao acaso outras 61 faturas, mas sem indicação de NIF, tendo-se verificado que a soma dos montantes foi de 1324 euros. Assume-se que o montante de faturas sem NIF também tem distribuição normal e que a variância das faturas com ou sem NIF é 50. Usando um teste adequado, comente a veracidade da afirmação "o montante médio das faturas com indicação de NIF é inferior ao das faturas sem NIF". (2.0)

3. Para inferir sobre a probabilidade θ de cancelamento de reservas num hotel por parte dos seus clientes, o gestor do hotel decidiu analisar os processos relativos a 230 reservas passadas, escolhidas ao acaso da sua vasta carteira, registando-se para cada reserva se a mesma tinha sido cancelada ou não.

(a) Mostre que o estimador do método dos momentos para θ é centrado. (1.0)

(b) De acordo com a desigualdade de Fréchet-Cramér-Rao, o estimador T da alínea anterior é o mais eficiente se: (1.0)

☐ $Var(T) = 0$ ☐ $Var(T) = \dfrac{\theta(1-\theta)}{230}$ ☐ $Var(T) < \dfrac{\theta(1-\theta)}{230}$ ☐ $Var(T) > \dfrac{\theta(1-\theta)}{230}$

(c) Sabendo que, das 230 reservas analisadas, 29 foram canceladas, construa um intervalo com nível de confiança aproximadamente 95% para θ. (1.5)

(d) Qual deveria ser a dimensão da amostra casual por forma a garantir que o “erro” cometido no intervalo da alínea anterior é inferior a 4%? (1.5)

TESTE II

1. Para cada afirmação assinale se esta é verdadeira **(V)** ou falsa **(F)**. (1.0)

	V	**F**
Se num teste de ajustamento que envolve 2 parâmetros desconhecidos e onde não foi necessário agrupar classes se tem, sob H_0, $Q \overset{a}{\sim} \chi^2_{(9)}$, onde Q é a estatística teste, então o número de classes é 11.		
No MRL a variável residual u é uma variável aleatória não observável porque depende dos coeficientes do modelo, que são desconhecidos.		
O teste de permanência de estrutura é um teste que permite detetar a presença de autocorrelação nos erros do modelo.		
Uma das hipóteses do MRL é a heterocedasticidade condicionada.		

2. Para testar se o desempenho profissional de determinada categoria de trabalhadores depende do género inquiriu-se uma amostra casual simples de 1000 trabalhadores tendo-se obtido: (2.0)

	género	
desempenho	masculino	feminino
baixo	170	120
médio	260	180
alto	170	100

Teste, ao nível de significância de 5%, a hipótese de independência entre o desempenho profissional e o género.

3. Para analisar o comportamento das vendas ($vendas$) das lojas de material de telecomunicação com base nas despesas em publicidade ($desp$) e na superfície da loja ($area$) em m^2, foi especificado o seguinte modelo

$$lvendas = \beta_1 + \beta_2 ldesp + \beta_3 area + \beta_4 cc + u$$

Os resultados obtidos encontram-se abaixo, onde o prefixo l representa o logaritmo natural e cc é uma variável artificial que assume o valor 1 se a loja se situa num centro comercial e o valor 0 caso contrário.

Regression Statistics	
Multiple R	0.6952
R Square	0.4833
Adjusted R Square	0.4744
Standard Error	0.2660
Observations	179

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	11.5849	3.8616	54.5590	0.0000
Residual	175	12.3863	0.0708		
Total	178	23.9712			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	9.9147	0.1395	71.0640	0.0000
ldesp	0.9620	0.2515	3.8256	0.0002
area	0.0062	0.0010	6.4936	0.0000
cc	0.1568	0.0649	2.4174	0.0167

(a) Interprete as estimativas de β_3 e de β_4 e analise a significância estatística da variável artificial, considerando um nível de significância de 5%. (1.5)

(b) Construa um intervalo de confiança a 95% para a elasticidade das vendas relativamente às despesas em publicidade e teste se é de admitir que um aumento de 2% nas despesas em publicidade origine um aumento, em média, de 3% nas vendas, mantendo constantes as restantes variáveis. (2.0)

(c) Interprete o coeficiente de determinação R^2 e teste a significância global do modelo. (1.5)

(d) Para testar a nulidade conjunta de β_3 e β_4 foi estimado o modelo $lvendas = \beta_1 + \beta_2 ldesp + u$, tendo-se obtido $VR = 16.959$, onde VR é a variação residual. Então, sendo F a estatística de teste adequada, tem-se (1.0)

- ☐ Sob H_0, $F \sim F_{(2,177)}$ e $F_{obs} = 32.67$
- ☐ Sob H_0, $F \sim F_{(2,176)}$ e $F_{obs} = 32.48$
- ☐ Sob H_0, $F \sim F_{(2,175)}$ e $F_{obs} = 32.30$
- ☐ Sob H_0, $F \sim F_{(2,174)}$ e $F_{obs} = 32.11$

(e) Com base na mesma amostra e definindo $ncc = 1 - cc$, foi estimado o modelo (1.0)

$$lvendas = \alpha_1 + \alpha_2 ldesp + \alpha_3 area + \alpha_4 ncc + u$$

Apresente, justificando, a equação ajustada desse modelo.

Continuação da questão...