

**Instituto Superior de Economia e Gestão**  
**Universidade de Lisboa**  
**Estatística II – Licenciatura em Gestão**  
 Exame de Época de Recurso – 28 de Janeiro de 2015  
**Parte Teórica**

Nome: \_\_\_\_\_ Nº: \_\_\_\_\_

**1. Perguntas de Verdadeiro/Falso [1.5 valores]**

Para cada afirmação assinale se esta é verdadeira (V) ou falsa (F). Uma resposta certa vale 0.3 e uma resposta errada penaliza em idêntico valor.

	V	F
Sejam $T_1$ e $T_2$ estimadores para um parâmetro $\theta$ . Se $Var(T_1) < Var(T_2)$ então $EQM(T_1) < EQM(T_2)$ .		
Se num teste para a média $\mu$ de uma população normal, $H_0: \mu = \mu_0$ contra $H_1: \mu = \mu_1$ , a potência do teste é igual a 0.90, pode concluir-se que a dimensão do teste é igual a 0.10.		
Sejam $(a, b)$ e $(c, d)$ intervalos de confiança a 90 % e 95% , respectivamente, para a média $\mu$ de uma população normal, obtidos pelo método habitual e com base na mesma amostra. Pode então concluir-se que $a > c$ e $b < d$ .		
Quando se acrescenta uma variável explicativa artificial num modelo de regressão linear, o coeficiente de determinação, $R^2$ , pode diminuir.		
Considere o modelo $y_t = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + u_t$ e os intervalos de confiança a 95 % para $y_0 = \beta_1 + \beta_2 c_2 + \beta_3 c_3 + u_0$ e para $E(y x_2 = c_2, x_3 = c_3)$ , obtidos com a mesma amostra. A amplitude do primeiro intervalo de confiança é superior à do segundo.		

**2. Perguntas de escolha múltipla [2.25 valores]**

Para cada pergunta assinale com **X** a alternativa correcta. Uma resposta certa vale 0.75 valores e uma resposta errada penaliza em 0.25 valores.

a) Considere o seguinte teste  $H_0: \lambda = 3$  contra  $H_1: \lambda > 3$  para o parâmetro  $\lambda$  de uma população exponencial de média  $E(X) = 1/\lambda$ . Com base numa amostra casual de dimensão  $n$  retirada dessa população, a região crítica uniformemente mais potente é da forma:

- $W = \{(x_1, x_2, \dots, x_n): \sum_{i=1}^n x_i > k\}$ .
- $W = \{(x_1, x_2, \dots, x_n): \sum_{i=1}^n x_i < k\}$ .
- $W = \{(x_1, x_2, \dots, x_n): \sum_{i=1}^n x_i < k_1 \vee \sum_{i=1}^n x_i > k_2\}$ .
- $W = \{(x_1, x_2, \dots, x_n): \sum_{i=1}^n x_i > 3n\}$ .

b) Num teste à bondade do ajustamento,  $H_0: p_j = p_{0j}$  contra  $H_1: p_j \neq p_{0j}$ , para algum  $j$ , onde  $Q$  é a estatística de teste sob  $H_0$ , e  $q_{obs}$  o valor observado de  $Q$ , o valor- $p$  é dado por:

- $P[Q \leq q_{obs}|H_0]$ .
- $P[Q \geq q_{obs}|H_0]$ .
- $2 \times P[Q \leq q_{obs}|H_0]$ .
- $2 \times P[Q \geq q_{obs}|H_0]$ .

c) Considere o modelo de regressão linear  $y_t = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + u_t$ . Se o coeficiente de correlação entre  $x_{t2}$  e  $x_{t3}$  é igual a 0.90, então

- os erros  $u_t$  são heterocedásticos.
- os erros  $u_t$  estão autocorrelacionados.
- o modelo deixa de ser linear.
- nenhuma das hipóteses do MRLC é afectada pela correlação elevada entre as variáveis explicativas.

**3. Perguntas de desenvolvimento [2.25 valores: a) 1valor; b) 1.25 valores]**

a) Seja  $(X_1, X_2, \dots, X_n)$  uma amostra casual de uma população normal de média  $\mu$  e variância  $\sigma^2$ . Estude a consistência do seguinte estimador para  $\mu$ :

$$T = \frac{1 + \sum_{i=1}^n X_i}{n}.$$

b) Considere o MRL  $y_t = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + u_t$  e a função de regressão ajustada  $\hat{y}_t = b_1 + b_2 x_{t2} + b_3 x_{t3}$ . Mostre que se verifica a igualdade  $\bar{y} = b_1 + b_2 \bar{x}_2 + b_3 \bar{x}_3$ , onde  $\bar{y}$ ,  $\bar{x}_2$  e  $\bar{x}_3$  são as médias amostrais de, respectivamente,  $y$ ,  $x_2$  e  $x_3$ .

**Instituto Superior de Economia e Gestão**  
**Universidade de Lisboa**  
**Estatística II – Licenciatura em Gestão**  
Exame de Época de Recurso – 28 de Janeiro de 2015

**Parte Prática**

Nome: \_\_\_\_\_ Nº : \_\_\_\_\_

**Espaço reservado para classificações**

1a. (20) \_\_\_\_      1c. (15) \_\_\_\_      2a. (15) \_\_\_\_      2c. (15) \_\_\_\_      2e. (15) \_\_\_\_

**Teórica:**

1b. (15) \_\_\_\_      1d. (20) \_\_\_\_      2b. (10) \_\_\_\_      2d. (15) \_\_\_\_

**Prática:**

**Em todos os testes de hipóteses que realizar, formule as hipóteses em teste, indique a estatística de teste e a sua distribuição. Para os intervalos de confiança proceda de forma semelhante para a variável fulcral. Considere sempre uma dimensão de teste de 0.05. Se necessitar de espaço para continuar qualquer resposta dispõe de uma página em branco no fim do enunciado.**

1. Um hipermercado recebeu um lote de pacotes de cereais de um novo fornecedor que garante que a quantidade média de gordura contida num pacote não ultrapassa 8 gramas. Para apresentar uma eventual reclamação ao fornecedor, o técnico do departamento de vendas recolheu uma amostra aleatória de 25 pacotes de cereais, tendo obtido uma quantidade média de gordura de 8.2 gramas. Admita que a quantidade de gordura contida num pacote de cereais do novo fornecedor segue uma distribuição normal de variância igual a 0.64.
  - a) Acha necessário reclamar? Justifique com base num teste adequado com dimensão de 5%.

b) Calcule a potência do teste quando a quantidade média de gordura contida num pacote de cereais é de 8.4 gramas.

c) Apresente uma estimativa para a percentagem de pacotes de cereais daquele fornecedor que contêm uma quantidade de gordura superior a 8.5 gramas (Nota: a estimativa de máxima verosimilhança para  $\mu$  é  $\bar{x}$ ).

- d) O gerente do hipermercado afirma que os clientes não manifestam preferência por qualquer dos três tipos de cereais (*clássico*, *com frutos vermelhos*, *com vitaminas e ferro*) da marca C&S à venda no hipermercado. Para analisar a afirmação do gerente, o técnico recolheu dados sobre 1500 consumidores de cereais da marca C&S, tendo concluído que 520 clientes escolheram os cereais *clássico* enquanto 450 optaram pelo cereais *com frutos vermelhos*. Teste a afirmação do gerente com base num teste adequado.

2. Para explicar as despesas das famílias em bens culturais (*dcult*) foi proposto o seguinte modelo:

$$ldcult_t = \beta_1 + \beta_2 lrend_t + \beta_3 educ_t + \beta_4 filhos_t + u_t$$

onde *ldcult* representa o logaritmo das despesas em bens culturais, *lrend* representa o logaritmo do rendimento médio da família, *educ* é o número médio de anos de escolaridade dos adultos da família e *filhos* é uma variável artificial que assume o valor 1 se a família tem filhos. Os resultados obtidos com base numa amostra de 320 observações encontram-se no Anexo

**Nota:** Admita que se verificam todas as hipóteses do MRLC para as questões das alíneas a), b), c) e d).

- a) Interprete a estimativa do parâmetro  $\beta_4$  e teste a significância estatística de  $\beta_4$ .

- b) Escreva o modelo ajustado para as famílias com filhos.

- c) Construa um intervalo de confiança a 95% para a elasticidade das despesas em bens culturais relativamente ao rendimento e com base nesse resultado, comente a seguinte afirmação: “É de admitir uma elasticidade unitária”.

d) Especifique um modelo que lhe permita testar  $H_0: \beta_2 = 2\beta_4$  contra  $H_1: \beta_2 \neq 2\beta_4$ . Admita que na estimação desse modelo obteve  $VR = 148.85$  (variação residual ou SSR). Que conclusão pode retirar?

e) A equação 2 é a regressão auxiliar de um teste. Qual? Efectue esse teste. Que conclusão pode retirar dos resultados obtidos?

## Anexo

**Equação 1:**  $ldcult_t = \beta_1 + \beta_2 lrend_t + \beta_3 educ_t + \beta_4 filhos_t + u_t$

<i>Regression Statistics</i>	
Multiple R	0.3902
R Square	0.1522
Adjusted R Square	0.1442
Standard Error	0.6848
Observations	320

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	3	26.6155	8.8718	18.9165	0.0000
Residual	316	148.2044	0.4690		
Total	319	174.8200			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	0.0736	0.1444	0.5099	0.6105
lrend	0.6336	0.1036	6.1160	0.0000
educ	0.0181	0.0052	3.4431	0.0007
filhos	0.2936	0.0810	3.6233	0.0003

**Equação 2:**  $\hat{u}_t^2 = \alpha_1 + \alpha_2 yfit_t + \alpha_3 yfit_t^2 + v_t$

(Nota:  $\hat{u}_t$  são os resíduos do modelo 1 e  $yfit$  são os valores ajustados da variável dependente do modelo 1)

<i>Regression Statistics</i>	
Multiple R	0.0486
R Square	0.0024
Adjusted R Square	-0.0039
Standard Error	0.8962
Observations	320

ANOVA					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	2	0.6034	0.3017	0.3756	0.6872
Residual	317	254.6322	0.8033		
Total	319	255.2356			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	0.4878	0.1151	4.2383	0.0000
YFIT	0.1356	0.4959	0.2735	0.7846
YFIT2	-0.2829	0.5198	-0.5443	0.5866