



ESTATÍSTICA II – Lic. Economia e Finanças
Época de Recurso – 29 de Janeiro de 2015 – Duração: 2h

Nome: _____ N.º: _____

<i>Espaço reservado a classificações</i>						
1.a)	1.b)	1.c)	2.a)	2.b)	2.c)	3.
4.a)	4.b)	4.c)	4.d)	4.e)	4.f)	4.g)
						Total

Perguntas de escolha múltipla: cada resposta certa vale 10 pontos; cada resposta errada vale -2.5 pontos; assinale a resposta escolhida com uma cruz no quadrado adequado. As cotações das restantes perguntas são indicadas no enunciado.

1. Para analisar a frequência de idas semanais ao ginásio numa certa cidade, efetuou-se um inquérito a uma amostra casual de 100 pessoas, cujos resultados foram os seguintes:

Nº Idas	0	1	2	≥ 3
Nº Pessoas	35	30	25	10

- a) [15] Teste a 5% se o número de idas semanais ao ginásio nesta cidade segue uma distribuição de Poisson de média 1. Que pode concluir?

- b) [20] Um estudo sobre uma segunda cidade recolheu informação sobre uma amostra casual de 200 pessoas, tendo-se observado que 115 delas frequentam o ginásio. Justificando através de um teste adequado, comente a veracidade da seguinte afirmação: “A proporção de cidadãos que frequentam o ginásio na primeira cidade é maior que na segunda”.

- c) Considere a seguinte afirmação: “Sempre que se observar uma diferença entre as proporções amostrais maior que 10 pontos percentuais, conclui-se que a proporção de cidadãos que frequentam o ginásio nas duas cidades é diferente”. Isto é equivalente a efectuar um teste com região crítica:

$W = \{(\bar{x}_1, \bar{x}_2) : \bar{x}_1 - \bar{x}_2 > 0.10\}$	
$W = \{z : z_{\text{obs}} > 0.10\}$	
$W = \{(\bar{x}_1, \bar{x}_2) : \bar{x}_1 - \bar{x}_2 > 0.10\}$	
Qualquer uma das anteriores, desde que $P[(\bar{X}_1, \bar{X}_2) \in W] = 0.05$	

2. Admita uma amostra casual de dimensão 4 de uma variável aleatória $X \sim N(\mu, 1)$.

- a) O limite superior de um intervalo de confiança para μ é de 3.288. Sabendo que da amostra se observou: $\bar{x} = 2$, indique o grau de confiança associado a este intervalo.

90%	
95%	
99%	
Não é possível obter o grau de confiança com a informação dada.	

- b) Considere agora o estimador para μ : $T = a_1X_1 + a_2X_2 + a_3X_3 + a_4X_4$, onde a_i ($i = 1, 2, 3, 4$) são constantes. Qual a restrição que as constantes a_i devem verificar para que o estimador T seja centrado?

$a_i = 1 \quad (i = 1, 2, 3, 4)$	
$a_1 + a_2 + a_3 + a_4 = 1$	
$a_1 = 0, a_2 = 1, a_3 = 1, a_4 = 0$	
$a_i < 0.25 \quad (i = 1, 2, 3, 4)$	

- c) [15] Considere agora uma amostra de dimensão $n > 3$ da mesma variável aleatória; estude a consistência do seguinte estimador para o parâmetro μ :

$$T = \frac{1}{n-1} \sum_{i=2}^{n-1} X_i$$

e compare-o com a média amostral em termos de eficiência.

3. [20] A percentagem de alunos com aprovação a Estatística II pode ser considerada uma variável aleatória X com a seguinte função densidade:

$$f(x) = \left(\frac{\lambda}{100^\lambda} \right) x^{\lambda-1} \quad (0 < x < 100; \lambda > 0)$$

Sabe-se também que: $E(X) = 100 \left(\frac{\lambda}{\lambda+1} \right)$.

Admita que nos três últimos semestres se registaram as seguintes percentagens de aprovação: (64, 67, 72). Obtenha a estimativa de máxima verosimilhança para a percentagem de aprovação média a Estatística II.

4. Com o objetivo de tentar identificar quais os factores determinantes dos gastos em cultura numa certa região, foi efetuado um inquérito a 500 consumidores entre os 25 e 40 anos. Começou-se por especificar o seguinte modelo econométrico:

$$\log(\text{gastos}) = \beta_0 + \beta_1 \log(\text{rend}) + \beta_2 \text{educ} + \beta_3 \text{idade} + \beta_4 \text{idade}^2 + u$$

onde *gastos* são os gastos em cultura, em euros; *rend* é o rendimento mensal do consumidor, em euros; e *educ* e *idade* são os anos de educação e idade do consumidor, respectivamente.

Os resultados da estimação OLS encontram-se no **Anexo**, bem como o de outra regressão. Se não houver indicação do contrário, assuma que se verificam as hipóteses clássicas do modelo de regressão linear.

- a) [15] Na **Equação 1**, interprete as estimativas dos coeficientes β_1 e β_2 , e teste a significância estatística dos anos de educação.

- b) Na **Equação 1**, o intervalo de confiança a 95% para a variação percentual estimada nos gastos em cultura por cada ano adicional de educação (*ceteris paribus*) é aproximadamente...

$(\hat{\beta}_2 \pm 1.96 \text{ se}(\hat{\beta}_2)) / 100$	
$\beta_2 \pm 1.96 \sigma / \sqrt{n}$	
$\hat{\beta}_2 \pm 1.96 \text{ se}(\hat{\beta}_2)$	
$100 \times (\hat{\beta}_2 \pm 1.96 \text{ se}(\hat{\beta}_2))$	

- c) Qual é aproximadamente a idade do consumidor, a partir da qual o efeito estimado sobre os gastos em cultura começa a ser positivo?

33 anos	
37 anos	
25 anos	
40 anos	

d) [20] Com base num teste adequado, diga se a idade do consumidor é estatisticamente relevante para explicar os gastos em cultura.

e) [15] Recorrendo à **Equação 2** e admitindo a normalidade do termo de erro, efectue uma previsão para o gasto médio mensal em cultura para os consumidores com 1500 euros de rendimento mensal e 15 anos de educação.

- f) [20] Admita que para averiguar a existência de heterocedasticidade na **Equação 2**, obtém o seguinte output do *EViews*:

Heteroskedasticity Test: Breusch-Pagan-Godfrey			
Obs*R-squared	18.71810	Prob. Chi-Square(2)	0.0001

Escreva a regressão auxiliar e efetue o respetivo teste. Que pode concluir sobre as propriedades estatísticas do estimador OLS aplicado na **Equação 2**?

- g) Qual das seguintes alternativas pode provocar inconsistência no estimador OLS?

Variância do termo de erro não constante.	
Inclusão de uma variável explicativa irrelevante.	
Não normalidade do termo de erro.	
Omissão de uma variável explicativa relevante.	

ANEXO

Equação 1

Dependent Variable: LOG(GASTOS)

Method: Least Squares

Included observations: 500

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	3.129708	1.407935	2.222907	0.0267
LOG(REND)	0.324694	0.032123	10.10795	0.0000
EDUC	0.018269	0.004979	3.669210	0.0003
IDADE	-0.114258	0.084419	-1.353453	0.1765
IDADE^2	0.001732	0.001270	1.364059	0.1732
R-squared	0.254253	Mean dependent var		3.976291
Adjusted R-squared	0.248227	S.D. dependent var		0.265042
S.E. of regression	0.229804	Akaike info criterion		-0.093231
Sum squared resid	26.14088	Schwarz criterion		-0.051085
Log likelihood	28.30770	Hannan-Quinn criter.		-0.076693
F-statistic	42.19104	Durbin-Watson stat		1.855937
Prob(F-statistic)	0.000000			

Equação 2

Dependent Variable: LOG(GASTOS)

Method: Least Squares

Included observations: 500

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.275433	0.225615	5.653138	0.0000
LOG(REND)	0.323633	0.031898	10.14601	0.0000
EDUC	0.017899	0.004960	3.608488	0.0003
R-squared	0.251363	Mean dependent var		3.976291
Adjusted R-squared	0.248350	S.D. dependent var		0.265042
S.E. of regression	0.229785	Akaike info criterion		-0.097363
Sum squared resid	26.24219	Schwarz criterion		-0.072075
Log likelihood	27.34066	Hannan-Quinn criter.		-0.087440
F-statistic	83.43655	Durbin-Watson stat		1.848197
Prob(F-statistic)	0.000000			