INSTITUTO SUPERIOR DE ECONOMIA E GESTÃO

UNIVERSIDADE DE LISBOA

ESTATÍSTICA II – LICENCIATURA EM GESTÃO

Exame de Época Normal – 12 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
Um estimador $\hat{\theta}$ é um estimador não enviesado de θ se $Var(\hat{\theta}) = Var(\theta)$		
Um intervalo de confiança para a proporção de uma população de Bernoulli, obtido pelo método habitual, é simétrico em torno da proporção amostral.		
Quanto mais afastada estiver a verdadeira média de uma população, μ, do valor μ_0 proposto no teste H_0: $\mu = \mu_0$, maior será a potência do teste.		
No modelo $y_t = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + u_t$, a variância da variável residual é observável.		
No modelo $y_t = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + u_t$, quanto maior for a variação do regressor x_{t2} na amostra, menor será a variância dos estimadores de mínimos quadrados de β_2, tudo o resto igual.		

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. Seja X_1, X_2 e X_3 uma amostra aleatória de uma população com média μ e variância σ^2. Sejam $T_1 = \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3$ e $T_2 = \frac{1}{6}X_1 + \frac{4}{6}X_2 + \frac{1}{6}X_3$ dois estimadores para μ. Então:

☐ T_1 é mais eficiente do que T_2;

☐ T_1 é menos eficiente do que T_2;

☐ A eficiência de T_1 é igual à de T_2;

☐ Nada se pode concluir sobre a eficiência de T_1 e T_2.

b. Suponha que de uma população normal X com média μ e variância σ^2 conhecida, se extraiu uma amostra aleatória de dimensão n, para testar as hipóteses $H_0: \mu = \mu_0$ contra $H_1: \mu > \mu_0$. Se a dimensão deste teste for α, a regra de decisão é:

☐ Rejeitar H_0 se $(\bar{x} - \mu_0)/\sqrt{\sigma^2/n} < -z_\alpha$.

☐ Rejeitar H_0 se $(\bar{x} - \mu_0)/\sqrt{\sigma^2/n} < z_\alpha$;

☐ Rejeitar H_0 se $(\bar{x} - \mu_0)/\sqrt{\sigma^2/n} > z_{\alpha/2}$;

☐ Rejeitar H_0 se $(\bar{x} - \mu_0)/\sqrt{\sigma^2/n} > z_\alpha$.

c. De acordo com o teorema de Gauss-Markov, quando satisfeitas as hipóteses apresentadas nas aulas, o estimador de mínimos quadrados dos coeficientes da regressão é:

☐ RED;

☐ BLACK;

☐ WHITE;

☐ BLUE.

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

a) Seja X uma variável aleatória com distribuição uniforme $U(0, \theta)$, com $\theta > 0$. Mostre que a média amostral, $\bar{X}$, é um estimador enviesado para θ, e encontre um estimador centrado para θ.

b) Explique em que situação existe heterocedasticidade. Que efeito tem a heterocedasticidade nos resultados de uma regressão pelo método dos mínimos quadrados?

INSTITUTO SUPERIOR DE ECONOMIA E GESTÃO

UNIVERSIDADE DE LISBOA

ESTATÍSTICA II – LICENCIATURA EM GESTÃO

Exame de Época Normal – 12 de Janeiro de 2015

PARTE PRÁTICA

Nome: ____ **Nº**___________

Espaço reservado para classificações				
1. (20)	2. (15)	3. (15)	4. (20)	**T:**
5a. (15)	5b. (15)	5c. (20)	5d. (20)	**P:**

ATENÇÃO! 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.

1. Considere uma variável aleatória X com função densidade de probabilidade

$$f(x) = \frac{x^{\alpha-1}e^{-\frac{x}{\beta}}}{\beta^{\alpha}\Gamma(\alpha)}, x > 0, \alpha, \beta > 0,$$

onde $\Gamma(\cdot)$ é a função gama. Tendo sido observada uma amostra $x_1, x_2, \dots, x_n$, e supondo que o parâmetro α é conhecido, encontre a estimativa de máxima verosimilhança do parâmetro β. Considere apenas a condição de primeira ordem.

2. Numa amostra aleatória de 160 alunos das licenciaturas do ISEG, 80 declararam pretender seguir um percurso profissional numa área quantitativa. Um intervalo de confiança para a proporção de alunos do ISEG que pretendem o mesmo é (0.4225; 0.5775). Indique, justificando, o grau de confiança deste intervalo.

3. Com o objectivo de investigar se no ISEG existem diferenças na assiduidade às aulas entre estudantes do sexo masculino e feminino, foram registados os tempos de presença nas aulas, ao longo de um semestre, de uma amostra aleatória de 20 estudantes do sexo masculino e 20 do sexo feminino. Foram obtidos os seguintes resultados:

Homens: $\sum_{i=1}^{20} x_{1i} = 2080$ horas, $\sum_{i=1}^{20} (x_{1i} - \bar{x}_1)^2 = 12844$ horas^2

Mulheres: $\sum_{i=1}^{20} x_{2i} = 2600$ horas, $\sum_{i=1}^{20} (x_{2i} - \bar{x}_2)^2 = 3211$ horas^2

Teste a hipótese de haver diferença entre sexos no que diz respeito ao tempo médio passado na sala de aulas, com uma dimensão de teste de 5%. Assuma que os tempos de presença nas aulas de estudantes do sexo masculino e feminino seguem distribuições normais com variância igual.

4. Numa amostra de 200 dias, o número de tentativas de intrusão nos sistemas informáticos de uma instituição do ensino superior teve a seguinte distribuição:

Número de tentativas por dia	0	1	2	3	4	5
Número de dias	44	67	46	26	12	5

Teste a hipótese, com uma dimensão de teste de 5%, de o número de tentativas de intrusão nos sistemas informáticos seguir uma distribuição de Poisson.

5. Um analista financeiro desenvolveu um modelo econométrico que pretende explicar o salário anual dos directores gerais de empresas em função do logaritmo do volume médio de vendas, em milhões de euros, nos 5 anos anteriores ($lvendas$), da rendibilidade média dos capitais próprios nos 5 anos anteriores em % (rcp), da rendibilidade média das acções nos 5 anos anteriores em % (ra), e de uma variável artificial ($servpub$), que assume o valor 1 se a empresa fornece serviços públicos e 0 caso contrário. Com base numa amostra aleatória de 209 empresas foi estimado o **Modelo 1**, que se encontra no Anexo, onde ($lsal$) representa o logaritmo do salário anual do director geral em milhares de euros.

a) Interprete a estimativa para o coeficiente da regressão da variável $servpub$ e teste a sua significância.

b) Teste a adequação global do modelo.

c) Uma empresa teve um volume médio de vendas de 10000 milhões de euros, uma rendibilidade média dos capitais próprios de 10% e uma rendibilidade média das acções de 13%. Além disso, esta empresa não fornece serviços públicos. Apresente um intervalo de previsão a 95% para o salário do director geral desta empresa.

d) O analista dividiu os dados em duas amostras: uma amostra contém 171 empresas com directores gerais que frequentaram o ensino superior; a outra amostra contém 38 empresas com directores gerais que frequentaram apenas o ensino secundário. Para cada uma destas amostras estimou uma regressão usando o Modelo 1. A variação residual dada pela primeira amostra foi 36.04, enquanto a da segunda amostra foi 6.45. Indique o que o analista pretende testar com estas regressões, efectue esse teste, com dimensão de 5%, e apresente as suas conclusões.

CONTINUAÇÃO DA QUESTÃO _____

ANEXO

Modelo 1: $lsal_t = \beta_1 + \beta_2 lvendas_t + \beta_3 rcp_t + \beta_4 ra_t + \beta_5 servpub_t + u_t$

Regression Statistics	
Multiple R	0.5829
R Square	0.3397
Adjusted R Square	0.3268
Standard Error	0.4647
Observations	209

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	22.6687	5.6672	26.2431	0.0000
Residual	204	44.0535	0.2159		
Total	208	66.7222			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	4.6283	0.3126	14.8060	0.0000
lvendas	0.2614	0.0343	7.6292	0.0000
rcp	0.0119	0.0042	2.8645	0.0046
ra	0.0003	0.0005	0.4804	0.6314
servpub	-0.3798	0.0904	-4.1989	0.0000

Modelo 2: $lsal_t = \beta_1 + \beta_2(lvendas_t - \log(10000)) + \beta_3(rcp_t - 10) + \beta_4(ra_t - 13) + \beta_5 servpub_t + u_t$

Regression Statistics	
Multiple R	0.5829
R Square	0.3397
Adjusted R Square	0.3268
Standard Error	0.4647
Observations	209

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	22.6687	5.6672	26.2431	0.0000
Residual	204	44.0535	0.2159		
Total	208	66.7222			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	7.1582	0.0543	131.8311	0.0000
lvendas – log(10000)	0.2614	0.0343	7.6292	0.0000
rcp – 10	0.0119	0.0042	2.8645	0.0046
ra – 13	0.0003	0.0005	0.4804	0.6314
servpub	-0.3798	0.0904	-4.1989	0.0000