

Instituto Superior de Economia e Gestão - UTL
Licenciaturas em Economia e Finanças
Estatística II, EN - 5 de Janeiro de 2011
Duração da prova: 2h

Instruções:

- Certifique-se de que o seu telemóvel está **desligado**.
 - **Formalize e fundamente** as suas respostas.
 - Durante o decorrer da prova não serão prestados quaisquer esclarecimentos. Se tiver alguma dúvida, apresente-a por escrito.
 - Caso não seja dito nada em contrário, utilize uma dimensão de 5% nos testes estatísticos que efectuar.
-

1 - Um ponto de venda de bilhetes da lotaria recolheu, através de uma amostra casual de 237 clientes, dados relativos às apostas, tendo-se observado uma média amostral de 23.37 Euros, com um desvio padrão corrigido de 11.68 Euros.

- a) Para testar se o valor da aposta pode ser considerado, em média, igual a 25 Euros, calcule um intervalo de confiança a 95%. Sem efectuar cálculos adicionais, diga se a sua conclusão poderia ser diferente se usasse um grau de confiança de 90%. Justifique.
- b) Seja μ o valor esperado da aposta na população (em Euros). Se se rejeitar a hipótese $H_0 : \mu = 25$ sempre que se observar uma média amostral superior a 26.5 ou inferior a 23.5, qual a probabilidade de cometer um erro de primeira espécie?

2 - Seleccionada uma amostra casual de 500 pessoas de várias regiões do país, perguntou-se a opinião sobre a escolha da “Pessoa do Ano” feita pela revista Time, tendo-se obtido os seguintes resultados:

		Região				
Opinião		Norte	Centro	Sul	Ilhas	Total
	Concorda	80	70	70	80	300
	Não concorda	70	80	30	20	200

- a) Será que as proporções de pessoas que **concordam** com a escolha **no Norte e no Centro** podem ser consideradas semelhantes? Responda com base num teste adequado, utilizando uma dimensão de 5%.
- b) Diga se a opinião pode ser considerada independente da região.

3 - Seja X uma variável aleatória contínua tal que $E(X) = k\theta$ e $\text{Var}(X) = k\theta^2$, onde k é uma constante positiva conhecida. Dada a amostra casual (X_1, \dots, X_n) , considere o seguinte estimador para θ :

$$T = \frac{\sum X_i}{kn}$$

Estude a consistência do estimador proposto. Sabendo que a Quantidade de Informação de Fisher para uma observação é dada por $I(\theta) = \frac{k}{\theta^2}$, diga se o estimador é o mais eficiente.

4 - No estudo dos factores que determinam o preço dos bilhetes de avião nos Estados Unidos, foram considerados diferentes modelos, com base nas seguintes variáveis:

FARE: preço do bilhete de ida (dólares);

LFARE: logaritmo natural de FARE;

DIST: distância entre origem e destino do voo (milhas);

LDIST: logaritmo natural de DIST;

LDISTSQ: quadrado de LDIST;

PASSEN: número médio de passageiros por dia na rota considerada (centenas);

BMKTSHR: quota de mercado da maior companhia a operar na rota considerada.

Tendo em conta os resultados da estimação dos modelos e das regressões auxiliares apresentados no Anexo, responda as seguintes questões:

- a) Considerando o modelo apresentado na **Equação 1**, interprete a estimativa do coeficiente da variável PASSEN. Será que pode considerar que o efeito do número de passageiros por dia sobre o preço do bilhete é nulo? Justifique através de um teste de hipótese adequado.
- b) Explique como pode testar a hipótese de homoscedasticidade no modelo da **Equação 1**, especificando o teste utilizado, as hipóteses nula e alternativa, a estatística de teste, a sua distribuição e a região crítica. Usando os resultados apresentados no Anexo, diga o que pode concluir sobre as propriedades dos estimadores e sobre os procedimentos efectuados na alínea anterior.
- c) Interprete o impacto estimado na **Equação 1** da distância no preço do bilhete, e compare com os resultados obtidos para este impacto na **Equação 2**. Através de um teste estatístico adequado, diga qual das formulações lhe parece mais correcta.
- d) Verifique se existe evidência de má especificação do modelo apresentado na **Equação 2** através de um teste adequado.
- e) Utilizando o modelo da **Equação 2**, faça a previsão do preço do bilhete para um voo com DIST=700, PASSEN=3000 e BMKTSHR=0.55.

ANEXO EN 05/01/2011**Equação 1**

Dependent Variable: LFARE

Included observations: 1149

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.423146	0.168663	14.36675	0.0000
LDIST	0.394479	0.021015	18.77150	0.0000
PASSEN	-0.004979	0.001486		
BMKTSHR	0.201667	0.063322	3.184807	0.0015
R-squared	0.359714	Mean dependent var		5.152683
Adjusted R-squared	0.358036	S.D. dependent var		0.417039
S.E. of regression	0.334143	Akaike info criterion		0.648977
Sum squared resid	127.8407	Schwarz criterion		0.666546
Log likelihood	-368.8373	Hannan-Quinn criter.		0.655609
F-statistic	214.4207	Prob(F-statistic)		0.000000

Ramsey RESET Test:

F-statistic	11.10076	Prob. F(2,1143)	0.0000
Log likelihood ratio	22.10408	Prob. Chi-Square(2)	0.0000

Breusch-Pagan-Godfrey Test:

F-statistic	61.63571	Prob. F(3,1145)	0.0000
Obs*R-squared	159.7542	Prob. Chi-Square(3)	0.0000
Scaled explained SS	109.2007	Prob. Chi-Square(3)	0.0000

Equação 2

Dependent Variable: LFARE

Included observations: 1149

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.029453	0.949740	7.401453	0.0000
LDIST	-1.022807	0.283467	-3.608201	0.0003
PASSEN	-0.005218	0.001435	-3.637137	0.0003
BMKTSHR	0.228381	0.061942	3.687004	0.0002
LDISTSQ	0.107560	0.021079	5.102795	0.0000
R-squared	0.377124	Mean dependent var		5.152683
Adjusted R-squared	0.374946	S.D. dependent var		0.417039
S.E. of regression	0.329712	Akaike info criterion		0.623149
Sum squared resid	124.3644	Schwarz criterion		0.645110
Log likelihood	-352.9993	Hannan-Quinn criter.		0.631440
F-statistic	173.1604	Prob(F-statistic)		0.000000

Equação 3

Dependent Variable: LFARE

Included observations: 1149

White Heteroskedasticity-Consistent Standard Errors & Covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	257.8080	251.6065	1.024648	0.3057
LDIST	-49.45452	48.79912	-1.013430	0.3111
PASSEN	-0.251085	0.247423	-1.014801	0.3104
BMKTSHR	11.00865	10.84118	1.015448	0.3101
LDISTSQ	5.197665	5.125398	1.014100	0.3107
FITTED^2	-9.090698	9.047003	-1.004830	0.3152
FITTED^3	0.581881	0.573670	1.014312	0.3106
R-squared	0.378333	Mean dependent var	5.152683	
Adjusted R-squared	0.375067	S.D. dependent var	0.417039	
S.E. of regression	0.329680	Akaike info criterion	0.624688	
Sum squared resid	124.1230	Schwarz criterion	0.655433	
Log likelihood	-351.8830	Hannan-Quinn criter.	0.636294	
F-statistic	115.8328	Prob(F-statistic)	0.000000	

Nota: a variável FITTED representa os valores ajustados da Equação 2.

Equação 4

Dependent Variable: FARE

Method: Least Squares

Date: 01/04/11 Time: 16:23

Sample: 1 4596 IF YEAR=2000

Included observations: 1149

Variable	Coefficient	Std. Error	t-Statistic	Prob.
EFIT	1.047892	0.009683	108.2190	0.0000
R-squared	0.375763	Mean dependent var	188.0235	
Adjusted R-squared	0.375763	S.D. dependent var	76.84232	
S.E. of regression	60.71210	Akaike info criterion	11.05103	
Sum squared resid	4231480.	Schwarz criterion	11.05543	
Log likelihood	-6347.818	Hannan-Quinn criter.	11.05269	

Nota: a variável EFIT representa o exponencial dos valores ajustados da Equação 2.