

Instituto Superior de Economia e Gestão - UTL
Licenciaturas em Economia e Finanças
Estatística II, ER - 26 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 - Seja X a variável aleatória que representa o número de comboios do metropolitano que chegam à estação do Rato. A v. a. X segue um processo de Poisson de taxa média de λ comboios por hora. Tendo-se observado uma amostra casual de 150 horas, obteve-se uma média amostral $\bar{x} = 12.72$.

- a) Efectuando um teste adequado, diga se se pode considerar que em média cheguem à estação do Rato mais que 12 comboios por hora.
- b) Escreva a função potência do teste da alínea anterior, e calcule o seu valor para $\lambda = 13$.
- c) Calcule uma estimativa da proporção de horas em que chegam 8 comboios.

2 - Um parque de merendas foi dividido em quatro zonas (Norte, Sul, Leste e Oeste). À saída foi inquirida uma amostra casula de 200 famílias sobre o local escolhido para almoçar, tendo-se obtidos os seguintes resultados:

Norte	Sul	Leste	Oeste	Total
37	49	62	52	200

Diga se se pode afirmar que as famílias são indiferentes à zona do parque. Justifique através de um teste adequado.

3 - Uma cadeia de cinemas abriu umas novas instalações (Cinema “Novo”), esperando cativar, em média, mais espectadores por espectáculo do que umas instalações mais antigas (Cinema “Antigo”). Seja X a variável aleatória que representa o número de espectadores por espectáculo. Tendo-se recolhido duas amostras casuais independentes do cinema “Antigo” e “Novo”, de 90 e 70 espectáculos, respectivamente, obtiveram-se os seguintes resultados:

Cinema	“Antigo”	“Novo”
\bar{x}	137.7	143.6
s'	7.2	12.8
n	90	70

Justificando através de um teste adequado, diga se se pode afirmar que o objectivo foi alcançado.

4 - Para estudar os rankings das faculdades de direito dos Estados Unidos foi estimado o seguinte modelo de regressão linear múltipla, com base numa amostra casual de 93 faculdades:

$$\text{RANK}_i = \beta_0 + \beta_1 \text{COST}_i + \beta_2 \text{SALARY}_i + \beta_3 \text{AGE}_i + \beta_4 \text{AGE}_i^2 + \beta_5 \text{STUDFAC}_i + u$$

onde as variáveis têm o seguinte significado:

RANK: ranking da faculdade;

COST: propinas anuais (milhares de dólares);

SALARY: salário médio anual de um recém licenciado (milhares de dólares);

AGE: idade da faculdade;

STUDFAC: ratio alunos/professores.

Tendo em conta os resultados da estimação do modelo na Equação 1 e das restantes regressões auxiliares apresentadas no Anexo, responda as seguintes questões:

- a) Supondo satisfeitas as hipóteses do modelo de regressão linear, teste ao nível de 5% a significância estatística individual do parâmetro β_2 e interprete a sua estimativa. Justificando através de um teste adequado, diga se se pode considerar que o efeito da idade é quadrático.
- b) Um investigador afirmou que a Equação 2 é preferível à Equação 1. Através de um teste de hipótese adequado, diga se concorda com a afirmação.

Responda às alíneas seguintes supondo que agora o modelo de referência è o da Equação 2:

$$\text{RANK}_i = \beta_0 + \beta_1 \text{COST}_i + \beta_2 \text{SALARY}_i + \beta_3 \text{AGE}_i + u$$

- c) Sabendo que na Equação 3 a estimativa do termo independente traduz a previsão pontual do ranking duma faculdade com propinas anuais de 13000 dólares, salário médio anual de um recém licenciado de 40000 dólares, e uma idade de 90 anos, indique como foi obtida esta equação e calcule a previsão por intervalo de confiança a 95% para o “ranking” médio das faculdades com estas características.
- d) Avalie a hipótese de que o impacto de SALARY no ranking é o dobro do impacto de COST. Apresente os cálculos e as conclusões a que chegou.
- e) Explique como obter uma regressão auxiliar para testar a presença de heterocedasticidade no modelo de referência, especificando o teste a utilizar, 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 nas alíneas anteriores.

ANEXO ER 26/01/2011

Equação 1:

Dependent Variable: RANK
Included observations: 93

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	248.7358	16.68105	14.91128	0.0000
COST	-1.684244	0.715377	-2.354345	0.0208
SALARY	-2.746598	0.289095	-9.500681	0.0000
AGE	-0.508807	0.213983	-2.377787	0.0196
AGESQ	0.001546	0.001070		
STUDFAC	-2.801444	2.330311	-1.202176	0.2326
R-squared	0.753720	Mean dependent var		77.34409
Adjusted R-squared	0.739566	S.D. dependent var		47.71097
S.E. of regression	24.34822	Akaike info criterion		9.285136
Sum squared resid	51576.72	Schwarz criterion		9.448529
Log likelihood	-425.7588	Hannan-Quinn criter.		9.351109
F-statistic	53.25126	Durbin-Watson stat		1.433303
Prob(F-statistic)	0.000000			

Equação 2:

Dependent Variable: RANK
Included observations: 93

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	224.7490	9.758612	23.03084	0.0000
COST	-1.501402	0.710927	-2.111893	0.0375
SALARY	-2.739744	0.290961	-9.416204	0.0000
AGE	-0.215330	0.073604	-2.925513	0.0044
R-squared	0.744489	Mean dependent var		77.34409
Adjusted R-squared	0.735876	S.D. dependent var		47.71097
S.E. of regression	24.52011	Akaike info criterion		9.278922
Sum squared resid	53509.98	Schwarz criterion		9.387851
Log likelihood	-427.4699	Hannan-Quinn criter.		9.322905
F-statistic	86.44032	Durbin-Watson stat		1.427590
Prob(F-statistic)	0.000000			

Equação 3:

Dependent Variable: RANK
Included observations: 93

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	76.26139	2.548615	29.92268	0.0000
X1	-1.501402	0.710927	-2.111893	0.0375
X2	-2.739744	0.290961	-9.416204	0.0000
X3	-0.215330	0.073604	-2.925513	0.0044
R-squared	0.744489	Mean dependent var		77.34409
Adjusted R-squared	0.735876	S.D. dependent var		47.71097
S.E. of regression	24.52011	Akaike info criterion		9.278922
Sum squared resid	53509.98	Schwarz criterion		9.387851
Log likelihood	-427.4699	Hannan-Quinn criter.		9.322905
F-statistic	86.44032	Durbin-Watson stat		1.427590
Prob(F-statistic)	0.000000			

Equação 4: (Nota: $Z = \text{COST} + 2 * \text{SALARY}$)

Dependent Variable: RANK
Included observations: 93

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	224.7490	9.758612	23.03084	0.0000
Z	-1.501402	0.710927	-2.111893	0.0375
SALARY	0.263061	1.591583	0.165282	0.8691
AGE	-0.215330	0.073604	-2.925513	0.0044
R-squared	0.744489	Mean dependent var		77.34409
Adjusted R-squared	0.735876	S.D. dependent var		47.71097
S.E. of regression	24.52011	Akaike info criterion		9.278922
Sum squared resid	53509.98	Schwarz criterion		9.387851
Log likelihood	-427.4699	Hannan-Quinn criter.		9.322905
F-statistic	86.44032	Durbin-Watson stat		1.427590
Prob(F-statistic)	0.000000			

Equação 5: (Nota: RES são os resíduos da Equação 2; RANKF são os valores ajustados da Equação 2).

Dependent Variable: RES^2
Included observations: 93

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	359.9145	155.4792	2.314872	0.0229
RANKF	0.743683	5.140219	0.144679	0.8853
RANKF^2	0.020623	0.042480	0.485467	0.6285
R-squared	0.038887	Mean dependent var		575.3761
Adjusted R-squared	0.017529	S.D. dependent var		669.5801
S.E. of regression	663.6856	Akaike info criterion		15.86522
Sum squared resid	39643073	Schwarz criterion		15.94692
Log likelihood	-734.7328	Hannan-Quinn criter.		15.89821
F-statistic	1.820725	Durbin-Watson stat		2.164127
Prob(F-statistic)	0.167821			