



ESTATÍSTICA II (Eco+Fin) - 2º Ano/1º sem.

Época de Recurso – 04.02.2020

1 hora (14 valores)

Nome: _____

Espaço reservado para classificações

1 a) (10)

1 b) (10)

1 c) (10)

2 a) (10)

2 b) (15)

2 c) (10)

3 (10)

4 a) (10)

4 b) i) (10)

4 b) ii) (10)

4 c) (15)

5 a) (10)

5 b) (10)

T:

NOTA: em todos os testes utilize um nível de significância de 5%.

Formalize e justifique as respostas a todas as questões.

1. Seja uma variável aleatória com distribuição uniforme $X \sim U(0, \theta)$, com $\theta > 0$.

a. Determine o estimador pelo método dos momentos para o parâmetro θ .

b. Mostre que a média da amostra é um estimador enviesado para θ e encontre um estimador centrado para o parâmetro θ .

c. De uma população normal, extraíram-se várias amostras de dimensão n . Comente a afirmação “ A variância da média da amostra é sempre igual à variância da população”.
Corrija se necessário

2. O controlo de qualidade de certos produtos, numa unidade fabril, é feito através da inspeção por amostragem desses produtos. Considere que o número de defeitos em cada produto é uma variável aleatória com distribuição de Poisson com média desconhecida. Numa amostra aleatória de 90 produtos registou-se um total de 9 defeitos. O intervalo de confiança a 99% para o número esperado de defeitos por produto para a amostra particular seleccionada é (0.0141, 0.1859).

a. A unidade fabril garante que “O número médio de defeitos por produto é igual a 0.05”. Comente a afirmação.

b. Determine o nível de confiança que reduz para 0.1 a amplitude do intervalo de confiança para a mesma amostra? Comente o resultado encontrado com base nos conceitos de precisão e confiança na estimação de parâmetros.

c. Explique **como** e **porque** varia a amplitude de um intervalo de confiança com a dimensão da amostra, mantendo-se constante o nível de confiança. Qual o efeito na precisão da estimativa?

3. Uma sondagem sobre as intenções de voto feita a 1000 eleitores nos EUA resultou na seguinte tabela de contingência:

		Sexo	
		Masculino	Feminino
Intenção de voto	Republicano	200	250
	Democrata	150	300
	Independente	50	50

Com base na informação que consta desta tabela e para um nível de significância de 5%, comente a independência entre o sexo do individuo e a sua inclinação partidária.

4. Por forma a estudar os determinantes das margens de lucro dos *franchisados* da empresa *XPTOStyle*, foi estimado o seguinte modelo (via Mínimos Quadrados Ordinários):

$$margem = \beta_0 + \beta_1 \log(vendas) + \beta_2 ncomp + \beta_3 nparc + \beta_4 ntemp + u$$

O *output* do modelo estimado:

Dependent Variable: MARGEM
Method: Least Squares
Included observations: 400

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	23.77211	3.932060	6.045713	0.0000
LOG(VENDAS)	1.703269	0.446501	3.814703	0.0002
NCOMP	0.195730	0.262909	0.744476	0.4570
NPARC	1.276986	0.371987	3.432877	0.0007
NTEMP	-1.473138	0.639404	-2.303925	0.0217
R-squared	0.083776	Mean dependent var		38.77423
Adjusted R-squared	0.074498	S.D. dependent var		5.218184
S.E. of regression	5.020050	Akaike info criterion		6.077178
Sum squared resid	9954.354	Schwarz criterion		6.127071
Log likelihood	-1210.436	Hannan-Quinn criter.		6.096936
F-statistic	9.029367	Durbin-Watson stat		1.859872
Prob(F-statistic)	0.000001			

Em que:

- *margem*, corresponde à margem de lucro do *franchisado i*;
- *vendas*, corresponde às vendas do *franchisado i* (em centenas de euros);
- *ncomp*, corresponde ao número de trabalhadores a tempo integral, do *franchisado i*;
- *nparc*, corresponde ao número de trabalhadores a tempo parcial, do *franchisado i*;
- *ntemp*, corresponde ao número de trabalhadores temporários (sazonais), do *franchisado i*.

(a) Interprete as estimativas para β_1 e para β_2 .

(b) Foi estimado um outro modelo:

$$\text{margem} = \beta_0 + \beta_1 \log(\text{vendas}) + \beta_2 \text{ncomp} + \beta_3 \text{nparc} + \beta_4 \text{ntemp} + \beta_5 \text{ncomp}^2 + \beta_6 \text{nparc}^2 + \beta_7 \text{ntemp}^2 + v$$

cujo output é dado no quadro seguinte:

Dependent Variable: MARGEM
Method: Least Squares
Included observations: 400

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	19.43422	4.748420	4.092776	0.0001
LOG(VENDAS)	1.740787	0.444215	3.918790	0.0001
NCOMP	-0.801716	0.930328	-0.861756	0.3893
NPARC	3.475134	0.841012	4.132088	0.0000
NTEMP	1.676619	2.537189	0.660818	0.5091
NCOMP^2	0.207267	0.158792	1.305272	0.1926
NPARC^2	-0.409514	0.142599	-2.871792	0.0043
NTEMP^2	-0.782505	0.597350	-1.309961	0.1910
R-squared	0.106448	Mean dependent var	38.77423	
Adjusted R-squared	0.090492	S.D. dependent var	5.218184	
S.E. of regression	4.976485	Akaike info criterion	6.067122	
Sum squared resid	9708.039	Schwarz criterion	6.146951	
Log likelihood	-1205.424	Hannan-Quinn criter.	6.098736	
F-statistic	6.671221	Durbin-Watson stat	1.862591	
Prob(F-statistic)	0.000000			

- i. Obtenha e interprete o ponto de viragem (caso exista) do regressor nparc^2 .

- ii. O que pode concluir acerca da significância conjunta dos regressores adicionados ao modelo?

(c) Considere o seguinte *output*:

Dependent Variable: RES^2
 Method: Least Squares
 Included observations: 400

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	197.7107	51.99378	3.802583	0.0002
LOG(VENDAS)	-16.49139	4.864026	-3.390481	0.0008
NCOMP	20.35428	10.18682	1.998100	0.0464
NPARC	-11.85814	9.208827	-1.287693	0.1986
NTEMP	-54.85223	27.78146	-1.974419	0.0490
NCOMP^2	-2.769698	1.738730	-1.592943	0.1120
NPARC^2	1.821772	1.561414	1.166745	0.2440
NTEMP^2	14.55501	6.540804	2.225263	0.0266
R-squared	0.053599	Mean dependent var		24.27010
Adjusted R-squared	0.036699	S.D. dependent var		55.51930
S.E. of regression	54.49103	Akaike info criterion		10.85375
Sum squared resid	1163955.	Schwarz criterion		10.93358
Log likelihood	-2162.749	Hannan-Quinn criter.		10.88536
F-statistic	3.171537	Durbin-Watson stat		2.018326
Prob(F-statistic)	0.002825			

Onde RES^2 corresponde ao quadrado dos resíduos do modelo estimado na alínea (b). Qual a informação que dele podemos retirar? Com base na conclusão anterior, discuta a validade dos resultados obtidos até aqui.

5. Considere o modelo

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i$$

em que u_i tem média nula e variância constante. Sabe-se ainda que uma amostra $\{(x_{1i}, x_{2i}, y_i): i = 1, 2, \dots, 1000\}$ foi retirada da população. Um aluno de Estatística II estimou a seguinte equação (via Mínimos Quadrados Ordinários):

$$\tilde{y}_i = \tilde{\beta}_0 + \tilde{\beta}_1 x_{1i}$$

Formalize e discuta o enviesamento e consistência (assim como a sua direção) do estimador dos Mínimos Quadrados Ordinários, nas seguintes situações:

(a) A variável x_{2i} está positivamente correlacionada com a variável x_{1i} , mas não tem qualquer efeito sobre y_i .

(b) A variável x_{2i} está negativamente correlacionada com a variável x_{1i} e tem um efeito positivo sobre y_i .