

Nome: _____ Nº _____

<i>Espaço reservado para classificações</i>		
1.a)	1.b)	1.c)
1.d)	1.e)	1.f)

Se necessitar de mais espaço, pode utilizar a última página do enunciado, indicando claramente a respectiva questão.
É expressamente proibido destacar as folhas do enunciado!

Boa Sorte!

1. Foi realizado um estudo econométrico com base nos principais indicadores de gestão de empresas *start-up* a operar em Portugal. O modelo empregue nesta análise é o seguinte:

$$\ln(VNEG_t) = \beta_1 + \beta_2 \ln(SERV_t) + \beta_3 \ln(PESS_t) + \beta_4 IMPOS_t + \beta_5 EMPR_t + u_t$$

Onde:

- $\ln(VNEG_t)$ – logaritmo do volume de negócios (em milhares de euros) da empresa t ;
- $\ln(SERV_t)$ – logaritmo das vendas de serviços (em milhares de euros) da empresa t ;
- $\ln(PESS_t)$ – logaritmo dos custos com o pessoal (em milhares de euros) da empresa t ;
- $IMPOS_t$ – impostos (em milhares de euros) pagos pela empresa t ;
- $EMPR_t$ – número de empregados da empresa t .

Em **ANEXO** encontram-se resultados da estimação de modelos de regressão linear obtidos com uma amostra de 139 empresas, e que deve utilizar para responder às questões que se seguem.

a) Interprete os coeficientes estimados dos regressores $\ln(SERV_t)$ e $EMPR_t$ na **Equação 1**. [1.0]

$b_2 = 0.3682774 \rightarrow$ *Ceteris paribus*, estima-se em média, que se as vendas de serviços aumentarem 1%, o volume de negócios aumenta aproximadamente 0.368%.

$b_5 = 0.0000079 \rightarrow$ *Ceteris paribus*, estima-se em média, que por cada empregado adicional, o volume de negócios aumenta aproximadamente 0.00079%.

- b) Teste a 5% a significância individual dos regressores $\ln(SERV_t)$ e $EMPR_t$ na **Equação 1**. [1.0]

Significância individual de $\ln(SERV)$:

$$H_0: \beta_2 = 0 \text{ vs } H_1: \beta_2 \neq 0$$

$$\mathbf{ET}: t_2 = \frac{b_2}{s_{b_2}} \sim t(134)$$

Do output obtém-se $p_{2,obs} = 0$. Porque $p_{2,obs} < \alpha = 0.05$, rejeita-se H_0 ao nível de 5%. Logo, o regressor $\ln(SERV)$ é estatisticamente significativo a 5%.

Significância individual de $EMPR$:

$$H_0: \beta_5 = 0 \text{ vs } H_1: \beta_5 \neq 0$$

$$\mathbf{ET}: t_5 = \frac{b_5}{s_{b_5}} \sim t(134)$$

Do output obtém-se $p_{5,obs} = 0.9712$. Porque $p_{5,obs} > \alpha = 0.05$, não se rejeita H_0 ao nível de 5%. Logo, o regressor $EMPR$ não é estatisticamente significativo a 5%.

- c) Interprete o coeficiente de determinação e teste a 5% a significância global da **Equação 1**. [2.0]

Coeficiente de Determinação:

$R^2 = 0.773524 \rightarrow$ o modelo explica 77.3524% da variação total do logaritmo do volume de negócios nesta amostra.

Significância Global:

$$H_0: \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0 \text{ vs } H_1: \exists \beta_j \neq 0 (j = 2, 3, 4, 5)$$

$$\mathbf{ET}: F = \frac{R^2/4}{(1 - R^2)/134} \sim F(4, 134)$$

Do output, obtém-se: $F_{obs} = 114.4185$.

$$W_{0.05} = \{F_{obs}: F_{obs} > 2.37\}$$

Porque $F_{obs} \in W_{0.05}$, rejeita-se H_0 ao nível de 5%. Logo, conclui-se que a Equação 1 é globalmente significativa.

- d) Na **Equação 1**, o que significa a hipótese $H_0: \beta_2 = \beta_3 = 0$? Sabendo que do modelo restrito se obteve $R^2 = 0.383$, efectue o teste de hipóteses correspondente (não se esquecendo de formalizar também a hipótese alternativa). Conclua utilizando um nível de 1% e interprete o resultado obtido. [2.0]

A hipótese $H_0: \beta_2 = \beta_3 = 0$ significa que os regressores $\ln(SERV)$ e $\ln(PESS)$ não são, no seu conjunto, significativos para explicar o volume de negócios das empresas.

$$H_0: \beta_2 = \beta_3 = 0 \text{ vs } H_1: \exists \beta_j \neq 0 (j = 2, 3)$$

$$\mathbf{ET}: F = \frac{(R^2 - R_0^2)/2}{(1 - R^2)/134} \sim F(2, 134)$$

$$W_{0.01} = \{F_{obs}: F_{obs} > 4.61\}$$

$$F_{obs} = \frac{(0.773524 - 0.383)/2}{(1 - 0.773524)/134} \approx 115.5315$$

Porque $F_{obs} \in W_{0.01}$, rejeita-se H_0 ao nível de 5%. Logo, conclui-se que os regressores $\ln(SERV)$ e $\ln(PESS)$ são conjuntamente significativos para explicar o volume de negócios das empresas.

- e) Usando a **Equação 2**, construa e interprete o intervalo de previsão com confiança de 90%, para a média do volume de negócios de empresas com vendas de serviços de 1808 milhares de euros, custos com pessoal de 8103 milhares de euros, impostos pagos de 250 milhares de euros e com 200 empregados. [2.0]

Previsão em média

IP a 90% para $\theta = E(\ln(VNEG_t) | SERV_t = 1808, PESS = 8103, IMP = 250, EMPR = 200)$

$$\mathbf{VF}: t_{\hat{\theta}} = \frac{\hat{\theta} - \theta}{s_{\hat{\theta}}} \sim t(134)$$

$$(\hat{\theta} \pm t_{0.005} \times s_{\hat{\theta}}) = (9.5854055 \pm 1.645 \times 0.1142598) = (9.39745, 9.77336)$$

IP a 90% para $E(VNEG_t | SERV_t = 1808, PESS = 8103, IMP = 250, EMPR = 200)$

$$(e^{9.39745}, e^{9.77336}) = (12057.595, 17559.669)$$

Logo, com uma confiança de 90% prevê-se que a média do volume de negócios para empresas com estas características se situa entre 12057.595 milhares de euros e 17559.669 milhares de euros.

- f) Qual o objectivo da estimação da **Equação 3** em **ANEXO**? Formalize o problema em termos das hipóteses a testar, estatística de teste e respectiva distribuição. O que pode concluir a 1%? [2.0]

O objectivo da Equação 3 é testar a presença de heterocedasticidade nos erros da Equação 1 (modelo em estudo) através do teste de White simplificado.

Assim, deve testar-se na regressão auxiliar de teste (Equação 3):

$$H_0: \alpha_2 = \alpha_3 = 0 \text{ vs } H_1: \exists \alpha_j \neq 0 (j = 2, 3)$$

$$\mathbf{ET}: W = nR_a^2 \overset{a}{\sim} \chi^2(2),$$

onde R_a^2 é o coeficiente de determinação da regressão auxiliar (Equação 3).

$$W_{0.01} = \{W_{obs}: W_{obs} > 9.210\}$$

$$W_{obs} = 139 \times 0.005772 = 0.802308$$

Porque $W_{obs} \notin W_{0.01}$, não se rejeita H_0 a 1%. Logo, existe evidência de homocedasticidade nos erros do modelo em estudo.

Continuação da questão: _____

ANEXO

Equação 1: $\ln(VNEG_t) = \beta_1 + \beta_2 \ln(SERV_t) + \beta_3 \ln(PESS_t) + \beta_4 IMPOS_t + \beta_5 EMPR_t + u_t$

<i>Regression Statistics</i>	
Multiple R	0.879502
R Square	0.773524
Adjusted R Square	0.766763
Standard Error	0.672571
Observations	139

ANOVA				
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Regression	4	207.0297	51.7574	114.4185
Residual	134	60.6152	0.4524	
Total	138	267.6449		

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	2.9142247	0.4172830	6.9838	0.0000
$\ln(SERV)$	0.3682774	0.0471631	7.8086	0.0000
$\ln(PESS)$	0.4299532	0.0725053	5.9300	0.0000
<i>IMPOS</i>	0.0001518	0.0000506	3.0017	0.0032
<i>EMPR</i>	0.0000079	0.0002171	0.0362	0.9712

Equação 2: O regressando é $\ln(VNEG_t)$.

<i>Regression Statistics</i>	
Multiple R	0.879502
R Square	0.773524
Adjusted R Square	0.766763
Standard Error	0.672571
Observations	139

ANOVA				
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Regression	4	207.0297	51.7574	114.4185
Residual	134	60.6152	0.4524	
Total	138	267.6449		

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	9.5854055	0.1142598	83.8913	0.0000
$\ln(SERV) - \ln(1808)$	0.3682774	0.0471631	7.8086	0.0000
$\ln(PESS) - \ln(8103)$	0.4299532	0.0725053	5.9300	0.0000
<i>IMPOS</i> - 250	0.0001518	0.0000506	3.0017	0.0032
<i>EMPR</i> - 200	0.0000079	0.0002171	0.0362	0.9712

Equação 3: $\hat{u}_t^2 = \alpha_1 + \alpha_2 \ln(\widehat{VNEG}_t) + \alpha_3 \ln(\widehat{VNEG}_t)^2 + v_t$

<i>Regression Statistics</i>	
Multiple R	0.075973
R Square	0.005772
Adjusted R Square	-0.008849
Standard Error	0.853878
Observations	139

ANOVA				
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>
Regression	2	0.575658573	0.2878	0.3948
Residual	136	99.15899786	0.7291	
Total	138	99.73465644		

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>
Intercept	1.593230241	3.018897724	0.5278	0.5985
$\ln(\widehat{VNEG}_t)$	-0.28811498	0.643748492	-0.4476	0.6552
$\ln(\widehat{VNEG}_t)^2$	0.017460542	0.033811041	0.5164	0.6064

Nota: \hat{u}_t são os resíduos e $\ln(\widehat{VNEG}_t)$ são os valores ajustados, ambos da **Equação 1**.