| LISBOA SCHOOL OF ECONOMICS & MANAGEMENT | **Estatística II**<br>Licenciatura em Gestão | 2º Semestre – 2018/2019<br>Época de Recurso<br>03/07/2019 |
|---|---|---|

Nome: ______________________________________________ Nº____________

| *Espaço reservado a classificações* | | | | | | |
|---|---|---|---|---|---|---|
| 1.a) | 1.b) | 2.a) | 2.b) | 3.a) | 3.b.i) | 3.b.ii) |
| 4.a) | 4.b) | 4.c) | 4.d) | 4.e) | 4.f) | |
| | | | | **TOTAL:** | | |

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. A nova *start-up* Xaquim desenvolveu um modelo de bateria para carros eléctricos que pensa apresentar num projecto à Tesla. O Elon Musk informou que só estaria interessado em tal modelo se tivesse um gasto energético inferior a 5.5 kWh, e então pediu à *start-up* que fosse feito um estudo em que se apresentasse um intervalo de confiança a 99% para a média do gasto energético da bateria. Recolhida uma amostra casual de 50 unidades, obteve-se: $\bar{x} = 5.61$ e $s' = 0.25$.

   a) Calcule e interprete o intervalo de confiança pedido pelo Elon Musk. Acha que o CEO da Tesla estará interessado na bateria da Xaquim? *[2.0]*

   b) De acordo com o intervalo de confiança anterior, o grau de confiança de 99% significa que…
   *[Resposta certa: 1 / Resposta errada: -0.25]*

| | |
|---|---|
| … a amplitude do intervalo de confiança é de 0.99. | |
| … a probabilidade do intervalo de confiança conter a verdadeira média é de 0.99. | |
| … a probabilidade da verdadeira média estar dentro do intervalo aleatório é de 0.99. | |
| … em 100 intervalos construídos, 99 são de confiança. | |

2. A Xaquim terminou um projecto de outro modelo de bateria que também pretende apresentar à Tesla, mas só o decidirá fazer se mais de 70% dos seus clientes estiverem satisfeitos. Tendo recolhido uma amostra casual de 100 clientes, 78 responderam estar satisfeitos enquanto os restantes afirmaram não estar satisfeitos.

   a) Qual será a conclusão da Xaquim, considerando um nível de 5%? *[2.0]*

   b) A dimensão de 5% do teste anterior significa que…
   *[Resposta certa: 1 / Resposta errada: -0.25]*

| | |
|---|---|
| … em 100 amostras casuais, o teste rejeita erradamente $H_0$ em 5. | |
| … a probabilidade de não rejeitar $H_0$ quando $H_0$ é verdadeira é de 0.05. | |
| … o respectivo valor-p é de 0.95. | |
| … a probabilidade de a dimensão da amostra ser grande é de 5%. | |

3. Assuma que o gasto energético de uma bateria é uma variável aleatória $X$ positiva com distribuição desconhecida de média $\mu > 0$ e variância $\sigma^2$. O Elon Musk enviou um *tweet* à Xaquim em que sugeria o seguinte estimador para a média do gasto energético:

$$\frac{1}{n}\sum_{i=1}^{n} \ln\left(\frac{X_i}{1+X_i}\right)$$

   a) Usando este estimador, em que intervalo de valores se encontra qualquer estimativa para $\mu$? Diga, justificando, se faz ou não sentido usar este estimador para $\mu$. *[1.0]*

b) Assuma que a Xaquim recolheu uma amostra casual de baterias de dimensão $n$ para obter uma estimativa pontual de $\mu$, mas ao analisar os dados, o novo estagiário perdeu inadvertidamente o valor do gasto energético da última observação. Neste cenário, a Xaquim irá optar por um dos seguintes estimadores:

$$T_1 = \frac{1}{n-1}\sum_{i=1}^{n-1} X_i \quad \text{e} \quad T_2 = \frac{1}{n}\sum_{i=1}^{n-1} X_i$$

Em relação ao estimador $T_1$, sabe-se que é centrado e consistente para $\mu$.

i) Mostre que o estimador $T_2$ é enviesado, mas consistente para $\mu$. *[2.0]*

ii) Sabe-se que $Var(T_2) < Var(T_1)$. Então…
*[Resposta certa: 1 / Resposta errada: -0.25]*

| | |
|---|---|
| … o estimador $T_2$ é mais eficiente que o estimador $T_1$. | |
| … nada se pode concluir quanto à eficiência. | |
| … o estimador $T_1$ é mais eficiente que o estimador $T_2$. | |

4. No seu papel de consultor estatístico, admita que foi contratado por um cliente para estudar o comportamento das vendas de um determinado produto em vários mercados com base nos gastos em publicidade. O modelo adoptado é o seguinte:

$$\ln(VENDAS) = \beta_1 + \beta_2 \ln(TV) + \beta_3 \ln(RADIO) + \beta_4 \ln(JORNAL) + u$$

Onde:

- $\ln(VENDAS)$ – logaritmo da quantidade vendida (milhares de unidades) do produto;
- $\ln(TV)$ – logaritmo do montante (milhares de euros) gasto em publicidade na televisão;
- $\ln(RADIO)$ – logaritmo do montante (milhares de euros) gasto em publicidade na rádio;
- $\ln(JORNAL)$ – logaritmo do montante (milhares de euros) gasto em publicidade nos jornais.

Todas as questões que se seguem referem-se sempre ao modelo inicial cuja estimação pode encontrar no **ANEXO** na **Equação 1**. No **ANEXO** também se encontram estimações de outras regressões que deve utilizar quando necessário como auxiliares de cálculo.

a) Interprete as estimativas dos parâmetros associados aos montantes gastos em publicidade na televisão e no jornal, e teste a significância individual de ambas as variáveis ao nível de 5%. *[2.0]*

b) Construa um intervalo de confiança a 95% para o parâmetro $\beta_3$ e interprete. *[1.0]*

c) Avalie a significância global da **Equação 1** ao nível de 5%. O que pode concluir? *[1.0]*

d) Teste a 5% a seguinte afirmação: "tudo o resto constante, o efeito sobre as vendas de um aumento de 1% nos gastos de publicidade na televisão é maior que o dobro do efeito de um aumento idêntico nos gastos em publicidade na rádio." O que pode concluir quanto à veracidade desta afirmação? (**Nota**: Utilize a **Equação 2** como auxiliar de cálculos.) *[2.0]*

e) Construa um intervalo de previsão a 99% para a quantidade vendida do produto num determinado mercado com os seguintes montantes gastos em publicidade: 150 mil euros em televisão, 25 mil euros em rádio e 30 mil euros em jornal. Interprete o resultado obtido. (**Nota**: Utilize a **Equação 3** como auxiliar de cálculos.) *[2.0]*

f) Para estudar a possível existência de comportamentos diferenciados na quantidade vendida do produto consoante a dimensão dos mercados, estimou-se separadamente dois modelos idênticos ao inicial: um com observações de grandes mercados do qual resultou $VR = 0.10861$; e outro com observações de pequenos mercados em que se obteve $VR = 1.05562$. Efectue um teste de hipóteses adequado ao nível de 5% e conclua. *[2.0]*

# ANEXO

**Equação 1:**

$$\ln(VENDAS) = \beta_1 + \beta_2 \ln(TV) + \beta_3 \ln(RADIO) + \beta_4 \ln(JORNAL) + u$$

| *Regression Statistics* | |
|---|---|
| Multiple R | 0.96084 |
| R Square | 0.92322 |
| Adjusted R Square | 0.92205 |
| Standard Error | 0.11569 |
| Observations | 200 |

ANOVA

| | *df* | *SS* | *MS* | *F* | *Significance F* |
|---|---|---|---|---|---|
| Regressão | 3 | 31.54247 | 10.51416 | 785.61330 | 0.00000 |
| Residual | 196 | 2.62314 | 0.01338 | | |
| Total | 199 | 34.16562 | | | |

| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* |
|---|---|---|---|---|
| Intercept | 0.44182 | 0.04838 | 9.13195 | 0.00000 |
| $\ln(TV)$ | 0.34935 | 0.00817 | 42.76287 | 0.00000 |
| $\ln(RADIO)$ | 0.15896 | 0.00767 | 20.72162 | 0.00000 |
| $\ln(JORNAL)$ | 0.01594 | 0.00859 | 1.85603 | 0.06495 |

**Equação 2:**

**Regressando:** $\ln(VENDAS)$

| *Regression Statistics* | |
|---|---|
| Multiple R | 0.96084 |
| R Square | 0.92322 |
| Adjusted R Square | 0.92205 |
| Standard Error | 0.11569 |
| Observations | 200 |

ANOVA

| | *df* | *SS* | *MS* | *F* | *Significance F* |
|---|---|---|---|---|---|
| Regression | 3 | 31.54247 | 10.51416 | 785.61330 | 0.00000 |
| Residual | 196 | 2.62314 | 0.01338 | | |
| Total | 199 | 34.16562 | | | |

| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* |
|---|---|---|---|---|
| Intercept | 0.44182 | 0.04838 | 9.13195 | 0.00000 |
| $\ln(TV)$ | 0.03144 | 0.01752 | 1.79488 | 0.07421 |
| $\ln(RADIO) + 2\ln(TV)$ | 0.15896 | 0.00767 | 20.72162 | 0.00000 |
| $\ln(JORNAL)$ | 0.01594 | 0.00859 | 1.85603 | 0.06495 |

**Equação 3:**

**Regressando:** $\ln(VENDAS)$

| *Regression Statistics* | |
|---|---|
| Multiple R | 0.96084 |
| R Square | 0.92322 |
| Adjusted R Square | 0.92205 |
| Standard Error | 0.11569 |
| Observations | 200 |

ANOVA

| | *df* | *SS* | *MS* | *F* | *Significance F* |
|---|---|---|---|---|---|
| Regression | 3 | 31.54247 | 10.51416 | 785.61330 | 0.00000 |
| Residual | 196 | 2.62314 | 0.01338 | | |
| Total | 199 | 34.16562 | | | |

| | *Coefficients* | *Standard Error* | *t Stat* | *P-value* |
|---|---|---|---|---|
| Intercept | 2.75816 | 0.00946 | 291.66454 | 0.00000 |
| ln(TV) – ln(150) | 0.34935 | 0.00817 | 42.76287 | 0.00000 |
| ln(RADIO) – ln(25) | 0.15896 | 0.00767 | 20.72162 | 0.00000 |
| ln(JORNAL) – ln(30) | 0.01594 | 0.00859 | 1.85603 | 0.06495 |