

Nome: _____ Número: _____

Notas: Os telemóveis devem estar desligados. O não cumprimento desta norma é motivo suficiente para anulação da prova. As perguntas de escolha múltipla têm cotação de **1** valor; **respostas erradas serão penalizadas em 0.2** valores. Não havendo informação em contrário, deverá utilizar um nível de significância de 5%. Pode usar a última página para continuar qualquer questão. Formalize devidamente todas as respostas.

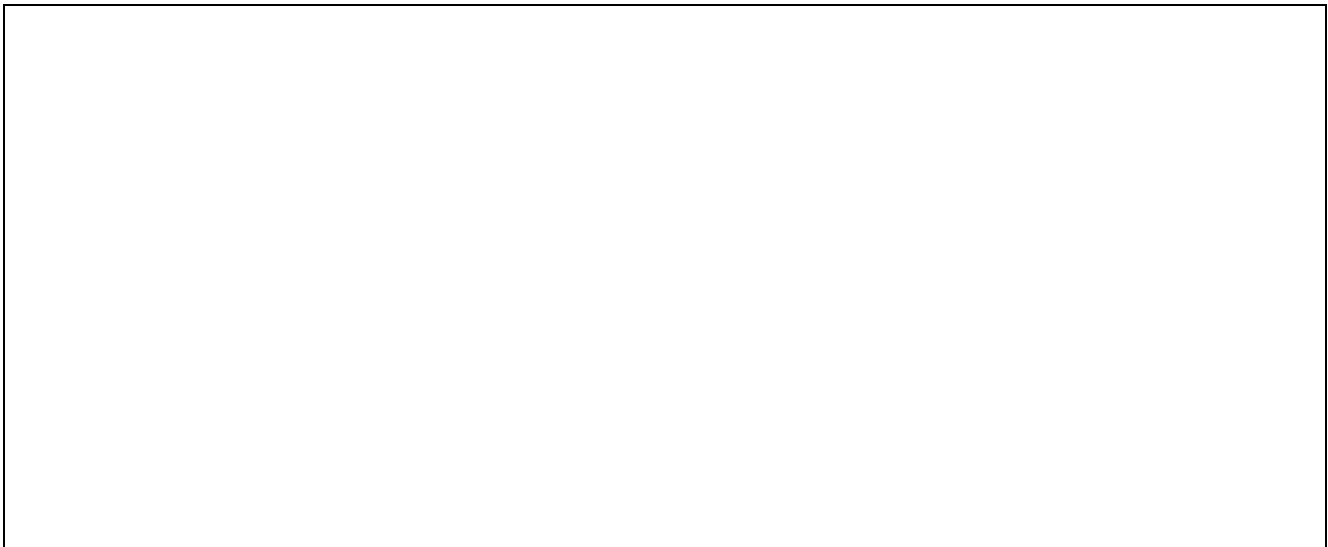
Espaço reservado para classificações

1. [2.0] Seja a variável aleatória X que representa o tempo de espera pelo primeiro cliente do dia numa agência bancária. Para testar se X tem distribuição exponencial, observou-se durante 100 dias o tempo que demorava a chegar o primeiro cliente à agência. Registou-se uma média de 10 minutos e a seguinte informação:

Tempo de espera (Minutos)	(0, 5]	(5, 10]	(10, 15]	(15, 20]	≥ 20
Frequência Observada	33	27	16	12	12
Frequência Esperada	A	B	14.475	8.779	13.534

Calcule as duas primeiras frequências esperadas em falta no quadro (**A** e **B**) e efetue o teste em questão. Que pode concluir com um nível de significância de 10%?

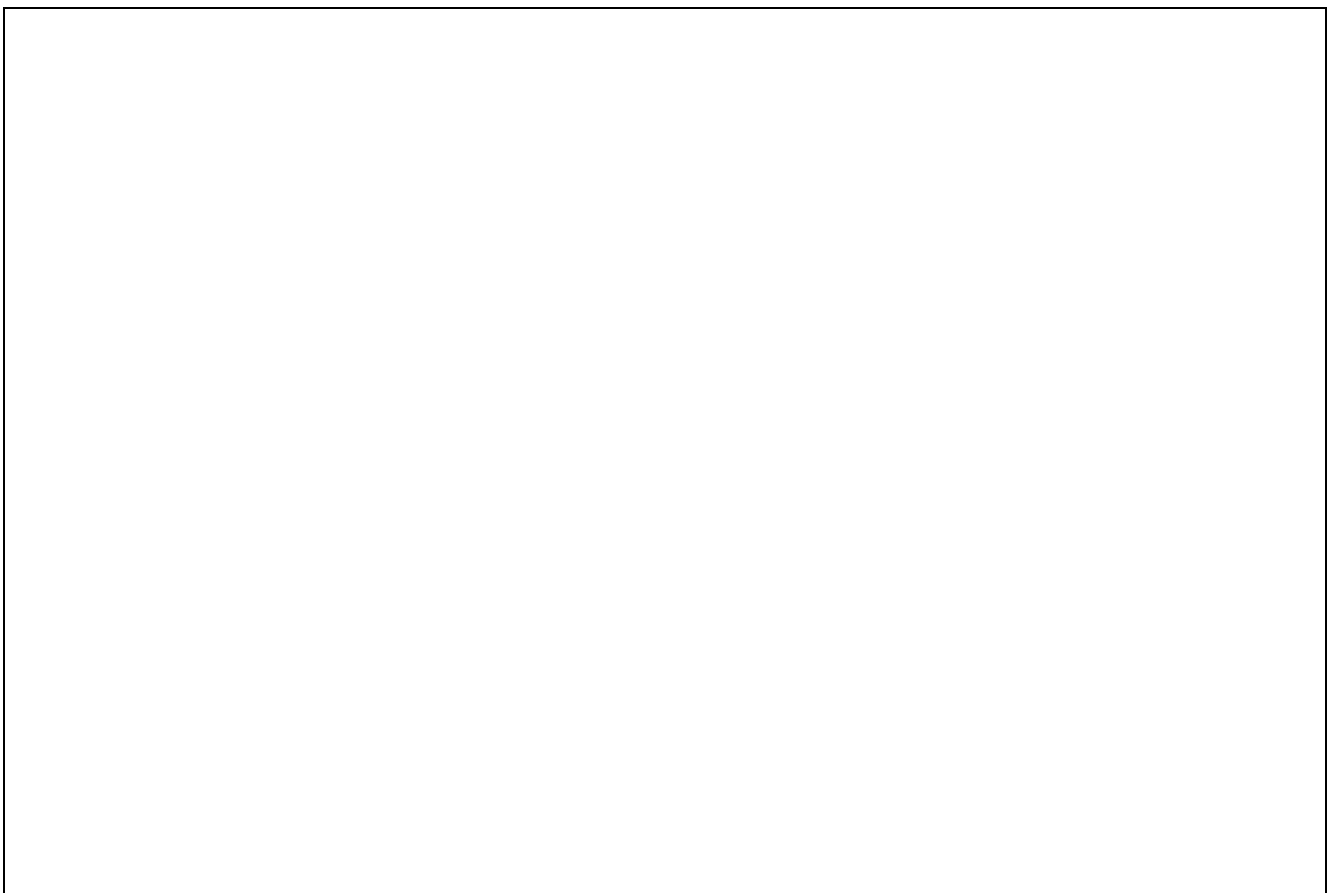
Continua na página seguinte



2. [1.5] Considere uma variável aleatória X com distribuição de Poisson de parâmetro λ . Admita que se retirou uma amostra aleatória de dimensão n . Considere o seguinte estimador para λ :

$$T_1 = \frac{4X_n - X_1 + \sum_{i=2}^{n-1} 2X_i}{2n}$$

Verifique se o estimador é centrado e consistente.



3. Suponha que se está a testar a independência entre o sexo de uma pessoa e a sua inclinação partidária dentro de m opções possíveis. Tendo-se observado uma amostra de dimensão n , indique qual das seguintes afirmações é verdadeira.

Os graus de liberdade da estatística de teste são dados por $n - m$.

A região de rejeição deste teste é bilateral porque a hipótese H_1 é a negação de H_0 .

- A inclinação partidária não é independente do sexo da pessoa se a estatística de teste for superior ao valor crítico da distribuição a 5%.
- As frequências esperadas são calculadas com base na distribuição do Qui-Quadrado.
- Nenhuma das respostas anteriores é verdadeira.

4. [1.5] Um restaurante de *fast food* pretende estimar a média de menus do tipo A consumidos diariamente através de um intervalo de confiança a 95%. Para tal observou o número de menus vendidos durante 50 dias obtendo os seguintes valores:

Mínimo	Máximo	Média	Desvio-padrão corrigido
5	45	20.3	12.5

Apresente a estimativa pretendida, indicando a variável fulcral que considerou.

5. [2.0] A gerência de uma empresa de produção de vinhos decidiu implementar um novo plano de marketing com vista a aumentar as suas encomendas. Antes de implementar o plano fez uma sondagem a 100 consumidores concluindo que apenas 33% se mostraram interessados em adquirir os seus vinhos. Depois do novo plano implementado fez nova sondagem entrevistando 120 consumidores concluindo que desta vez 40% pretendia adquirir os vinhos em questão. Poderá concluir-se que a diferença é estatisticamente significativa para se considerar que o plano de marketing foi eficaz?

Continua na página seguinte

6. Se um teste de hipótese simples contra hipótese simples tem potência 0.8 isto significa que,

- Quando H_1 é verdadeira, a probabilidade de H_0 não ser rejeitada é 0.2.
- Quando H_1 é verdadeira, a probabilidade de H_0 ser rejeitada é 0.2.
- Quando H_0 é verdadeira, a probabilidade de H_0 ser rejeitada é 0.8.
- Quando H_0 é verdadeira, a probabilidade de H_0 não ser rejeitada é 0.8.
- Nenhuma das afirmações anteriores é verdadeira.

7. Seja $(1, 2)$ o intervalo de confiança a 95% para μ obtido com uma amostra de dimensão n numa população normal com variância conhecida. Então,

- A probabilidade de $\mu = 1.5$ é exatamente 0.95.
- A probabilidade de μ pertencer ao intervalo $(1, 2)$ é 0.95.
- Dada a mesma amostra, o intervalo $(1.25, 1.75)$ para μ tem confiança superior a 0.95.
- A probabilidade da variável fulcral habitual pertencer ao intervalo $(-1.96, 1.96)$ é exatamente 0.95.
- Nenhuma das afirmações anteriores é verdadeira.

8. Considere o modelo de regressão linear que verifica as hipóteses de Gauss-Markov, $y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + u_i$, e os respetivos resíduos OLS, \hat{u}_i .

- O resíduo \hat{u}_i corresponde à observação da variável aleatória u_i na amostra particular observada.
- A variável residual u_i não é observável apesar de existir uma amostra para $y_i, x_{i1}, \dots, x_{ik}$.
- Como existe termo independente no modelo a soma dos quadrados dos resíduos é igual a zero.
- Embora as variáveis u_i sejam observáveis é preferível considerar antes os resíduos \hat{u}_i porque são obtidos substituindo os β_j pelo estimador OLS que é centrado e mais eficiente.
- Nenhuma das afirmações anteriores é verdadeira.

9. Com o intuito de estudar o salário dos gestores empresariais, foi recolhida uma amostra de 177 gestores, tendo-se estimado o seguinte modelo:

$$lsal_i = \beta_0 + \beta_1 lucros_i + \beta_2 lvendas_i + \beta_3 exper_i + \beta_4 exper2_i + u_i$$

Onde:

- *sal* – Salário do gestor, em milhares de euros;
- *lsal* – Logaritmo do salário;
- *lucros* – Lucros da empresa em milhões de euros;
- *vendas* – Vendas da empresa em milhões de euros;
- *lvendas* – Logaritmo das vendas;
- *exper* – Número de anos como gestor da empresa;
- *exper2* – Quadrado de *exper*.

No **Anexo** figuram os resultados obtidos na estimação deste modelo.

a) [2.0] Interprete as estimativas obtidas para β_1 e β_2 e teste a sua significância estatística individual.

b) **[2.0]** Formalize o efeito no salário de um gestor quando a respectiva experiência aumenta de 2 para 3 anos e estime este efeito. Teste a respectiva significância estatística.

c) **[1.5]** Será razoável afirmar que o salário do gestor não depende da sua experiência (EXPER e EXPER2)? Justifique através de um teste estatístico.

d) [1.5] Qual o objetivo da regressão incluída na equação 4 do anexo? Que pode concluir? Nota: a variável FIT corresponde aos valores ajustados para *lsal* obtidos com a equação 1.

10. Assinale a afirmação verdadeira.

- Se se verificarem as hipóteses de Gauss-Markov o estimador OLS dos coeficientes β_j não é igual ao estimador do método dos momentos.
- O estimador OLS tem como hipótese de que o modelo é linear, isto é, a variável dependente y tem necessariamente uma relação linear com as variáveis explicativas.
- O facto das variáveis explicativas estarem muito correlacionadas entre si não afeta a variância do estimador OLS mas sim a sua média.
- Existe heterocedasticidade quando a variância das variáveis residuais for muito grande.
- Nenhuma das afirmações anteriores é verdadeira.

11. Assinale a afirmação **FALSA**.

- O coeficiente de determinação, R^2 , aumenta sempre que diminui a soma dos quadrados dos resíduos para uma dada variável dependente.
- Adicionando uma variável que é irrelevante ao modelo pode diminuir o \bar{R}^2 .
- Com heterocedasticidade o R^2 mantém a mesma interpretação válida.
- O R^2 não deve ser usado para comparar modelos em que o número de variáveis explicativas é diferente.
- Se o R^2 é muito elevado então todas as variáveis explicativas são estatisticamente relevantes.

Continuação da questão _____

ANEXO

Equação 1

Dependent Variable: LSAL

Method: Least Squares

Included observations: 177

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.904169	0.239714	20.45845	0.0000
LUCROS	0.000184	0.000117	1.565510	0.1193
LVENDAS	0.196364	0.033142	5.924888	0.0000
EXPER	0.045375	0.014219	3.191118	0.0017
EXPER2	-0.001230	0.000478	-2.571800	0.0110
R-squared	0.333595	Mean dependent var		6.582848
Adjusted R-squared	0.318098	S.D. dependent var		0.606059
S.E. of regression	0.500468	Akaike info criterion		1.481296
Sum squared resid	43.08054	Schwarz criterion		1.571018
F-statistic	21.52536	Prob(F-statistic)		0.000000

Equação 2

Dependent Variable: LSAL

Method: Least Squares

Included observations: 177

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	5.145093	0.234732	21.91904	0.0000
LUCROS	0.000179	0.000120	1.483440	0.1398
LVENDAS	0.193699	0.033999	5.697170	0.0000
R-squared	0.289839	Mean dependent var		6.582848
Adjusted R-squared	0.281676	S.D. dependent var		0.606059
S.E. of regression	0.513659	Akaike info criterion		1.522291
Sum squared resid	45.90921	Schwarz criterion		1.576124
F-statistic	35.50747	Prob(F-statistic)		0.000000

Equação 3

Dependent Variable: LSAL

Method: Least Squares

Included observations: 177

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.904169	0.239714	20.45845	0.0000
LUCROS	0.000184	0.000117	1.565510	0.1193
LVENDAS	0.196364	0.033142	5.924888	0.0000
EXPER	0.040455	0.012463	3.245974	0.0014
EXPER2-4*EXPER	-0.001230	0.000478	-2.571800	0.0110
R-squared	0.333595	Mean dependent var		6.582848
Adjusted R-squared	0.318098	S.D. dependent var		0.606059
S.E. of regression	0.500468	Akaike info criterion		1.481296
Sum squared resid	43.08054	Schwarz criterion		1.571018
F-statistic	21.52536	Prob(F-statistic)		0.000000

Equação 4

Dependent Variable: LSAL

Method: Least Squares

Included observations: 177

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.993602	6.650713	0.750837	0.4538
LUCROS	0.000195	0.000874	0.223507	0.8234
LVENDAS	0.206805	0.776660	0.266275	0.7903
EXPER	0.047833	0.183221	0.261066	0.7944
EXPER2	-0.001296	0.004973	-0.260679	0.7947
FIT^2	-0.004126	0.306642	-0.013456	0.9893
R-squared	0.333596	Mean dependent var		6.582848
Adjusted R-squared	0.314111	S.D. dependent var		0.606059
S.E. of regression	0.501929	Akaike info criterion		1.492594
Sum squared resid	43.08050	Schwarz criterion		1.600260
Log likelihood	-126.0946	Hannan-Quinn criter.		1.536259
F-statistic	17.12023	Prob(F-statistic)		0.000000