LICENCIATURA EM GESTÃO
ESTATÍSTICA II

EXAME | 1º SEMESTRE | 2023-2024

ÉPOCA NORMAL | 22 DE JANEIRO DE 2024

Duração: 2 Horas

NOME (legível): ____

N.º de Aluno: ____

Note Bem:

- Leia atentamente todas as questões e todos os dados.
- Indique todos os cálculos necessários para a resolução de cada exercício na folha de exame, exceto se esse cálculo depender de outros já realizados.
- Cada uma das perguntas de escolha múltipla tem só uma resposta correta. Indique-a na folha de exame.
- Não é necessário justificar as respostas às perguntas de escolha múltipla.
- Respostas incorretas às perguntas de escolha múltipla não são descontadas na classificação do exame.
- Na resolução do Exame, pode alterar a ordem das perguntas.
- Pode utilizar a máquina de calcular básica e o formulário.
- O formulário e as tabelas encontram-se em anexo.
- O enunciado da prova fica com os docentes.

---

1. O número de automóveis que passam em cada período de um minuto num determinado ponto de uma autoestrada é uma variável aleatória com distribuição de Poisson de valor médio λ. Tendo por base a contabilização do número de automóveis que passam nesse ponto da autoestrada em cada um de 100 intervalos de um minuto, selecionados ao acaso:

1.1. Deduza o estimador de máxima verosimilhança (EMV) de λ. [2.0]

1.2. Será que o EMV de λ é centrado? [1.0]

1.3. Determine o erro quadrático médio do EMV de λ. [1.0]

----

2. Um vendedor de uma marca de baterias, usadas em *tablets*, mediu os tempos (em horas) de autonomia de 16 dessas baterias, x_i, i = 1, 2, …, 16, tendo obtido os seguintes dados: $\sum_{i=1}^{16} x_i = 79.5$ e $s'^2 = 0.0423$ (variância amostral corrigida). Admitindo que o tempo de autonomia dessas baterias tem distribuição normal:

2.1. Obtenha um intervalo de confiança (IC) a 95%, para o tempo médio de autonomia das baterias em questão. [2.0]

2.2. A amplitude de um intervalo de confiança a 95% para o valor médio (desvio padrão populacional desconhecido) de uma variável com distribuição normal não é influenciada por: [1.0]

A) A dimensão da amostra.
B) O nível/grau de confiança.
C) O desvio padrão da amostra.
D) A média da amostra.
E) O nível de significância 0.05.

2.3. Teste ao nível de significância de 5% se a variância do tempo de autonomia das baterias em questão é superior a 0.06. [2.0]

2.4. Num teste estatístico, sabe-se que: [1.0]

A) A região de rejeição não depende do nível de significância.
B) O valor-p não depende do nível de significância.
C) A probabilidade de erro tipo I é igual à potência do teste.
D) A probabilidade de erro tipo II é igual à potência do teste.
E) Se o valor-p é superior ao nível de significância (α), então rejeita-se a hipótese nula para α.

----

3. O registo, efetuado ao longo de um conjunto de várias semanas, de faltas ao trabalho dos funcionários de uma empresa conduziu aos seguintes resultados:

Dias	Segunda	Terça	Quarta	Quinta	Sexta
Nº de Faltas	23	17	14	20	26

3.1. Teste a hipótese de as faltas ao trabalho dos funcionários da empresa se distribuírem uniformemente pelos 5 dias úteis da semana para o nível de significância de 10%. [2.0]

3.2. Uma das condições de aplicabilidade do teste de hipóteses referido na alínea anterior é: [1.0]

A) A variância tem que ser homogénea nos grupos de estudo.
B) Os valores esperados têm que ter uma distribuição normal.
C) Pelo menos 80% dos valores observados têm que ser superiores a 5.
D) Pelo menos 80% dos valores esperados têm que ser superiores a 5.
E) A dimensão da amostra tem que ser superior a 20.

----

4. Este estudo pretende avaliar o impacto de alguns fatores que explicam a distância percorrida pelas famílias durante o período de férias. O modelo de regressão linear de referência é o seguinte:

$$\boldsymbol{dist = \beta_1 + \beta_2 rend + \beta_3 idade + \beta_4 criancas + u}$$

onde:

- $\boldsymbol{dist}$ **− distância percorrida nas férias, em quilómetros;**
- $\boldsymbol{rend}$ **− rendimento familiar anual, em milhares de euros;**
- $\boldsymbol{idade}$ **− idade média dos adultos no agregado familiar, em anos;**
- $\boldsymbol{criancas = \begin{cases} 1, & \text{se existem crianças no agregado familiar} \\ 0, & \text{caso contrário} \end{cases}}$
- $\boldsymbol{u}$ **− erro**

O *output* do *EXCEL* com o modelo estimado é o seguinte:

Regression Statistics	
Multiple R	0.607
R Square	0.369
Adjusted R Square	0.359
Standard Error	681.426
Observations	194

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	3	51585572.745	17195190.915	37.031	0.000
Residual	190	88224848.246	464341.307		
Total	193	139810420.991			

	Coefficients	*Standard Error*	*t Stat*	*p-value*
Intercept	-431.113	260.136	-1.657	0.099
rend	24.724	2.729	9.059	0.000
idade	17.383	5.543	3.136	0.002
crianças	-188.638	115.240	-1.637	0.103

4.1. Interprete os efeitos estimados do rendimento e da existência de crianças sobre a distância percorrida nas férias. [0.5]

4.2. O que pode concluir quanto à significância estatística da variável binária? [1.0]

A) É significativa para 5%.
B) É significativa para 10%.
C) É significativa para 5%, mas não para 10%.
D) É significativa para 10%, mas não para 5%.
E) Não é significativa para 10%.

4.3. Construa um intervalo de confiança a 95% para o parâmetro β_1 e interprete o resultado. [2.0]

4.4. O valor do coeficiente de determinação permite concluir que: [1.0]

A) Os regressores explicam mais de metade da variabilidade amostral da distância.
B) Os regressores explicam metade da variabilidade amostral da distância.
C) Fica por explicar 63.1% da variabilidade amostral da distância.
D) A probabilidade de o modelo estimado estar correto é de 0.369.
E) A probabilidade de o modelo estimado estar incorreto é de 0.369.

4.5. Teste a significância global do modelo ao nível de 1% [2.0]

4.6. A primeira família na amostra tem um rendimento anual de 41 mil euros, uma idade média dos adultos de 26 anos e não tem crianças. Obtenha o valor ajustado para esta observação e interprete. [0.5]