



### Exame de Recorrência de Métodos Estatísticos

Data: 09/07/2003

Duração: 2 horas

Nome: \_\_\_\_\_ N.º: \_\_\_\_\_

Curso: \_\_\_\_\_

Regime: \_\_\_\_\_ Número de folhas suplementares entregues pelo aluno: \_\_\_\_\_

Declaro que desisto \_\_\_\_\_

Esta prova consiste em 5 questões de resposta aberta e as respectivas cotações encontram-se na tabela que se segue.

Questão	1 a)	1 b)	1 c)	2 a)	2 b)	3 a)	3 b)	3 c)	4 a)	4 b)	5
Cotação	2	1	2	2	2	2,5	1,5	1,5	1,5	2	2
Classif.											

1. O passado dia 23 de Maio foi assinalado pela estreia europeia da sequência “*Matrix Reloaded*”.

O Gustavo, grande entusiasta de filmes de ficção, não quer perder a hipótese de voltar a ver *Neo, Trinity e Morpheus* no grande ecrã. Quando, em 1999, quis assistir à primeira parte desta saga, “*Matrix*”, o Gustavo foi obrigado a fazer 4 incursões às bilheteiras até conseguir um bilhete.

Considerando que a probabilidade de conseguir bilhete em cada ida ao cinema é  $\frac{1}{4}$ ,

a) Indique qual a distribuição da variável aleatória  $X =$  “número de idas à bilheteira necessárias para conseguir bilhete” e escreva a expressão da função massa de probabilidade,  $f(x)$ .

b) Com base na alínea anterior, calcule a probabilidade de o Gustavo conseguir o bilhete logo na primeira vez que se dirige à bilheteira.

E o Gustavo lá conseguiu o bilhete para o filme. Infelizmente, não foi à primeira nem à segunda: teve que se dirigir à bilheteira 5 vezes para o conseguir! Achou então por bem estimar o parâmetro da distribuição da variável aleatória  $X$ . Para isso, perguntou a vários amigos quantas tentativas de compra tiveram que fazer até conseguir o tal bilhete e obteve uma média  $\bar{x} = 5,5$ .

c) Utilizando o método dos momentos, obtenha uma estimativa para o parâmetro da distribuição considerada nas alíneas anteriores.

2. O Sr. Leandro, gerente dos cinemas em Aveiro, espera que as receitas de “*Matrix Reloaded*” superem as do primeiro filme. Após 31 dias de exibição, o Sr. Leandro admitiu reunir as condições necessárias para determinar um intervalo de confiança para o valor médio da receita diária de “*Matrix Reloaded*”. Paralelamente, recorreu a um teste de comparação de médias baseado nas receitas obtidas nos primeiros 31 dias de exibição de ambos os filmes (“*Matrix*” em 1999 e “*Matrix Reloaded*” em 2003). Admitindo que as duas amostras disponíveis são provenientes de populações normais, o Sr. Leandro obteve os resultados seguintes:

**Group Statistics**

	ano	N	Mean	Std. Deviation	Std. Error Mean
Receitas (milhares de euros)	1999	31	80,4802	4,82354	,86633
	2003	31	84,6731	7,30783	1,31253

**Independent Samples Test**

		t-test for Equality of Means		
		t	df	Sig. (2-tailed)
Receitas (milhares de euros)	Equal variances assumed	-2,666	60	,010

a) De acordo com os resultados do teste de hipóteses, diga a que níveis de significância há evidências para afirmar que as receitas diárias de “*Matrix Reloaded*” ultrapassam as anteriormente conseguidas com o primeiro filme.

b) Teste, ao nível de significância 10% as seguintes hipóteses sobre a variância  $\sigma^2$  das receitas diárias do segundo filme:

$$H_0: \sigma^2 = 65 \text{ vs } H_1: \sigma^2 < 65$$

Nome: \_\_\_\_\_ N.º: \_\_\_\_\_

3. Apesar dos resultados até à data serem bastante satisfatórios, o Sr. Leandro decidiu investigar se há ou não preferência por algumas sessões. Assim durante o mesmo período de tempo (31 dias), registou as receitas obtidas nas 4 sessões (15:00, 17:45, 21:30 e 0:15). Para averiguar se as receitas diárias diferem significativamente de sessão para sessão, procedeu-se à análise de variância (ANOVA).

a) Complete a tabela obtida:

**ANOVA**

receitas (milhares de euros)

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1863,700	_____	_____	_____	,000
Within Groups	138,483	_____	_____		
Total	_____	_____			

Uma análise mais detalhada foi levada a cabo através do método de Bonferroni de comparações múltiplas.

**Multiple Comparisons**

Dependent Variable: receitas (milhares de euros)

Bonferroni

(I) Sessões	(J) Sessões	Mean Difference (I-J)	Std. Error	Sig.
Sessão das 15:00	sessão das 17:45	-,4146	,27286	,788
	sessão das 21:30	-,6943	,27286	,073
	sessão das 00:15	8,5653*	,27286	,000
sessão das 17:45	Sessão das 15:00	,4146	,27286	,788
	sessão das 21:30	-,2797	,27286	1,000
	sessão das 00:15	8,9799*	,27286	,000
sessão das 21:30	Sessão das 15:00	,6943	,27286	,073
	sessão das 17:45	,2797	,27286	1,000
	sessão das 00:15	9,2596*	,27286	,000
sessão das 00:15	Sessão das 15:00	-8,5653*	,27286	,000
	sessão das 17:45	-8,9799*	,27286	,000
	sessão das 21:30	-9,2596*	,27286	,000

\*. The mean difference is significant at the .05 level.

b) Com base nos resultados apresentados na tabela, identifique os pares de médias que podem ser considerados diferentes ao nível de significância de 5%.

c) A aplicação do método LSD (método sem qualquer correcção ao nível de significância total) permitiria alterar a sua resposta à alínea anterior? Justifique.

4. O Sr. Leandro está a ponderar sobre a continuidade da última sessão (0:15). Toma então a decisão de suspender a última sessão por tempo indeterminado no caso de a audiência não chegar a rondar os 32 espectadores. Em 7 dias, seleccionados ao acaso, o número de espectadores que assistiram à última sessão foi:

Dia	1	2	3	4	5	6	7
n.º de espectadores	24	38	26	27	19	25	33

a) De entre os testes abaixo apresentados que foram aplicados à amostra referida,

**One-Sample Kolmogorov-Smirnov Test**

	número de espectadores na sessão das 0:15
N	7
Kolmogorov-Smirnov Z	,639
Exact Sig. (2-tailed)	,727

**One-Sample Test**

	Test Value = 32		
	t	df	Sig. (2-tailed)
número de espectadores na sessão das 0:15	-1,938	6	,101

**Wilcoxon Signed Ranks Test**

	número esperado de espectadores (32) - número de espectadores na sessão das 0:15
Z	-1,609 <sup>a</sup>
Exact Sig. (2-tailed)	,125

a. Based on negative ranks.

**Paired Samples Test**

		t	df	Sig. (2-tailed)
Pair 1	número de espectadores na sessão das 0:15 - número esperado de espectadores (32)	-1,938	6	,101

**Sign Test**

	número esperado de espectadores (32) - número de espectadores na sessão das 0:15
Exact Sig. (2-tailed)	,453 <sup>a</sup>

a. Binomial distribution used.

indique, justificando, qual considera ser o mais adequado. (Observação: A variável “número esperado de espectadores (32)” corresponde a uma variável de valor constante igual a 32.)

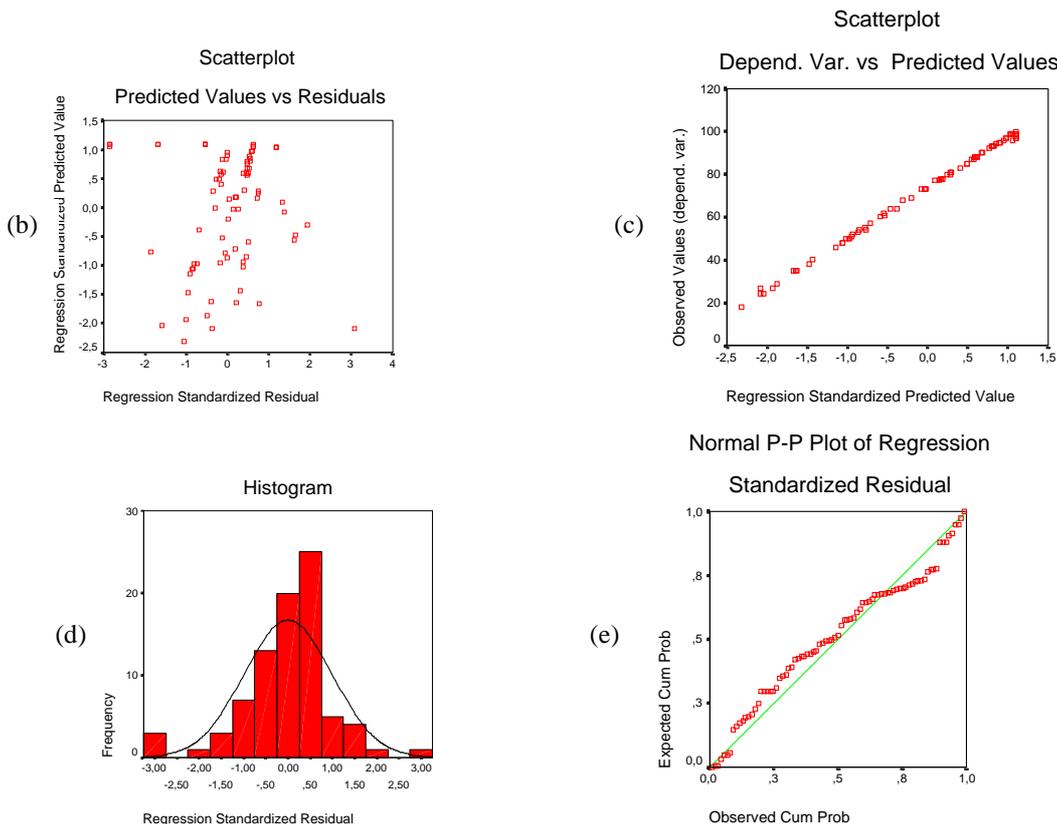
- b) Soube-se entretanto que o número  $X$  de espectadores da última sessão num dia escolhido ao acaso é uma variável aleatória com valor médio 28 e desvio padrão 7,2. Ao tomar conhecimento desta informação, o Sr. Leandro alterou a sua regra de decisão. Até nova ordem, a última sessão só será suspensa se, em 31 dias seleccionados ao acaso, o número total de espectadores for inferior a 900. Recorrendo ao Teorema do Limite Central, calcule a probabilidade de a sessão ser suspensa.

5. Apesar de tudo, o Sr. Leandro continua apreensivo... Está em causa a rentabilidade da sessão que depende, obviamente, de diversas variáveis: número de espectadores, número de funcionários necessários, electricidade dispendida, direitos de exibição, etc... Identificadas as variáveis de interesse, pretendeu-se estudar como é que elas influenciam a rentabilidade. Optou-se por ajustar um modelo de regressão linear, o qual permitiu derivar a tabela e construir os gráficos que se seguem:

**Model Summary**

(a)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,999	,999	,999	,862



Quais os gráficos e/ou tabelas que utilizaria para:

- validar os pressupostos do modelo? \_\_\_\_\_
- avaliar a qualidade do ajustamento? \_\_\_\_\_

Independentemente da rentabilidade ou não desta sessão, o Sr. Leandro pode dormir descansado: é que a estreia de “*Matrix Revolutions*” está marcada para 7 de Novembro e, como ninguém quer perder a batalha final entre Homens e Máquinas, o filme promete. E o Gustavo lá estará!