



Exame Especial de Métodos Estatísticos

Departamento de Matemática - Universidade de Aveiro

4 de Dezembro de 2003

Duração: 2h 30 min

NOME: _____

N^o mecan.: _____ Curso: _____ Turma: _____

[] Declaro que desisto: _____ Classificação: _____

Esta prova consiste em 5 questões de resposta aberta. As cotações são as seguintes:

Ques.	1a	1b	1c	2a	2b	3a	3b	3ci	3cii	3diii	4a	4b	4c	4d	4e	5a	5b	5c	6di	5dii
Cot.	0.8	1.2	1.5	1.5	2	1	.6	.8	.6	.5	.5	.9	1	.8	1	1	2	1.5	.5	1.5

Algumas questões da prova têm espaços para preencher com a forma de . Nestas questões não devem ser apresentados os cálculos efectuados.

Resolva as restantes questões justificando convenientemente os passos fundamentais.

Na consulta de tabelas não é necessário proceder a interpolação, basta utilizar os valores da tabela mais próximos dos pretendidos.

Entre 12 de Junho e 4 de Julho de 2004, Portugal será o centro do futebol europeu com a realização da fase final do UEFA EURO 2004.

1. Como é de conhecimento público, os formulários de candidatura aos bilhetes do torneio já estão disponíveis. Existem duas modalidades na aquisição de bilhetes. A primeira modalidade contempla a compra de bilhetes para jogos individuais; a segunda modalidade, denominada *Follow My Team*, permite seguir uma selecção durante o torneio. O Tiago candidatou-se a 30 bilhetes para jogos individuais e a 20 bilhetes *Follow My Team*. Suponha que a probabilidade de conseguir qualquer um destes bilhetes é 0.1.

Complete os espaços em branco:

- (a) A variável aleatória X que representa o número de bilhetes *Follow My Team* a atribuir ao Tiago tem distribuição de parâmetro(s) .
- (b) A variável aleatória que representa o número total de bilhetes a atribuir ao Tiago tem valor médio e segue uma distribuição de parâmetro(s) .

(c) Utilizando o Teorema do Limite Central, calcule a probabilidade aproximada do Tiago conseguir arranjar mais de 10 bilhetes, no total.

2. O Programa de Parceiros é de extrema importância para a validade do UEFA EURO 2004. Numa fase inicial deste programa, a organização previa celebrar contratos publicitários com grandes multinacionais. Admita que os montantes resultantes destes contratos podem ser considerados como provenientes de uma população Normal com variância conhecida, $\sigma^2 = 0.25$.

(a) Quantos contratos se devem celebrar por forma a se obter um intervalo de confiança, a 95%, para a média com amplitude 0.1.

(b) Suponha agora que se desconfia que a variância não é 0.25. Até ao momento já se celebraram 101 contratos que permitem contabilizar, em média, 0.9 milhões de euros, e o respectivo desvio padrão corrigido é 0.35. Teste, através do cálculo da região crítica e ao nível de significância de 1%, a hipótese de a variância da população ser inferior a 0.25.

NOME: _____ N^o.: _____

3. A 24 de Julho de 1999, 34310 pessoas juntaram-se no relvado do Estádio Nacional unidos por uma única causa: a realização do Euro 2004 em Portugal. Formaram então um logotipo humano que passou a figurar do *Guinness Book of Records*. Mediram-se os tempos, em segundos, entre chegadas consecutivas ao estádio dos primeiros 50 participantes, tendo-se posteriormente construído o seguinte diagrama de caule-e-folhas:

Tempos entre Chegadas dos Participantes

Frequency	Stem & Leaf
15,00	0 . 000001111111111
10,00	0 . 222233333
8,00	0 . 44444455
5,00	0 . 66777
4,00	0 . 8899
3,00	1 . 111
3,00	1 . 333
2,00	1 . 45
Stem width:	1,00
Each leaf:	1 case(s)

- (a) Se tivesse que escolher uma distribuição para modelar os tempos entre chegadas consecutivas, qual escolheria? Justifique.

(b) A mediana da amostra é e o máximo é .

- (c) As 34310 pessoas que se juntaram no relvado foram chegando ao Estádio Nacional ao longo de várias horas. Durante duas horas registou-se o número de pessoas que entravam no estádio ao longo de cada minuto. Efectuou-se um teste de ajustamento de Kolmogorov-Smirnov para averiguar se o número de pessoas que entravam no estádio em cada minuto podia ser modelado por uma distribuição de Poisson, e obtiveram-se os seguintes resultados:

One-Sample Kolmogorov-Smirnov Test

		CHEGADAS
N		120
Poisson Parameter ^{a,b}	Mean	99,8500
Most Extreme Differences	Absolute	,041
	Positive	,041
	Negative	-,036
Kolmogorov-Smirnov Z		,451
Asymp. Sig. (2-tailed)		,987

- a. Test distribution is Poisson.
- b. Calculated from data.

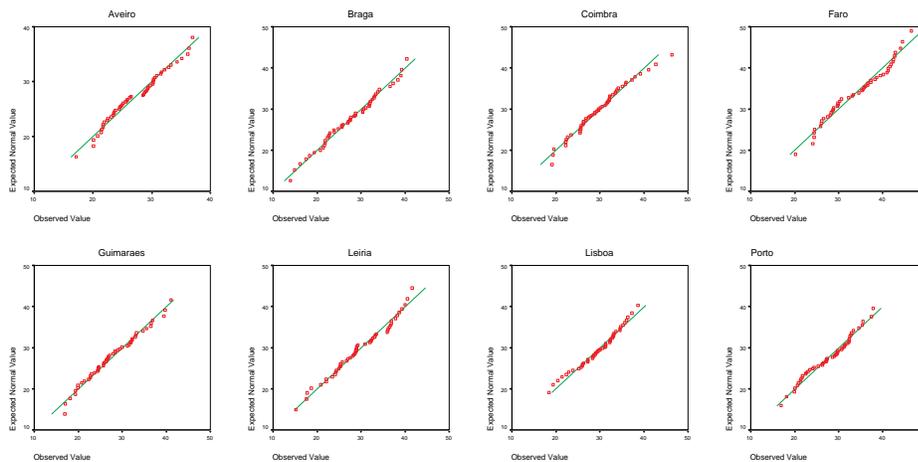
O que pode concluir aos níveis usuais de significância (1, 5 e 10%)?

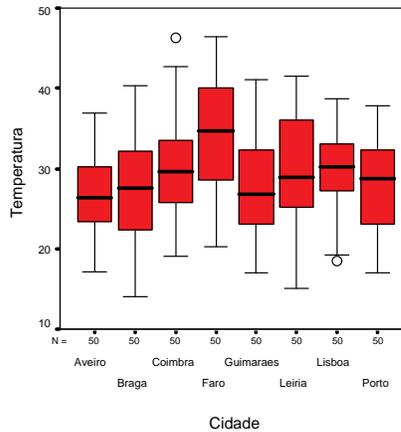
(d) Sabendo que a média da amostra considerada é $\bar{x} = 99.85$ obtinha, a partir do método dos momentos, uma estimativa para o parâmetro λ da distribuição de Poisson.

(e) Durante as duas horas de observação chegaram efectivamente ao estádio participantes.

4. As cidades de torneio foram escolhidas de acordo com vários critérios, entre eles o de terem um bom clima.
 Recolheram-se as temperaturas médias diurnas durante 50 dias de Verão escolhidos ao acaso para cada uma das 8 cidades e construíram-se os seguintes gráficos:

QQ-plots normais:





(a) Identifique o último gráfico.

Efectuaram-se também análises de variância (ANOVA) paramétrica e não-paramétrica tendo-se obtido os seguintes resultados:

ANOVA paramétrica:

ANOVA

Temperatura

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1329,858	_____	189,980	5,459	,000
Within Groups	13641,163	_____	34,799		
Total	14971,021	_____			

ANOVA não paramétrica:

Ranks

	Cidade	N	Mean Rank
Temperatura	Aveiro	50	163,74
	Braga	50	173,67
	Coimbra	50	213,76
	Faro	50	264,69
	Guimaraes	50	175,09
	Leiria	50	213,95
	Lisboa	50	219,38
	Porto	50	179,72
Total		400	

Test Statistics^{a,b}

	Temperatura
Chi-Square	29,858
df	7
Asymp. Sig.	,000

a. Kruskal Wallis Test
b. Grouping Variable: Cidade

(b) Complete a tabela de ANOVA.

(c) Diga, justificando, qual das análises utilizaria para testar se as cidades têm temperaturas semelhantes.

(d) Conclua, aos níveis usuais de significância (1, 5 e 10%), se todas as cidades escolhidas têm temperaturas semelhantes. Justifique.

(e) Foram efectuadas comparações múltiplas testando todos os pares de localizações sem qualquer correcção ao nível de significância do total das 16 comparações. Ao nível de 0.1, apenas os pares que envolvem a cidade de Faro acusam diferenças significativas, tendo os *p-values* dos referidos testes sido os seguintes:

Avei.-Faro	Braga-Faro	Coimb.-Faro	Guim.-Faro	Leiria-Faro	Lisb.-Faro	Porto-Faro
0.000	0.000	0.009	0.000	0.0061	0.0061	0.000

Aplicando o método de Bonferroni, diga quais os pares que acusam diferenças significativas, ao nível de significância global 0.1. Justifique.

5. A construção do estádio municipal de Aveiro, um dos estádios do torneio, foi adjudicada no Verão de 2001. Os trabalhos de terraplanagem foram adjudicados à firma Construtores Casais pelo valor de 450 mil euros. Ao longo de 40 semanas efecturam-se registos semanais de despesas da firma (em milhares de euros) bem como de outras variáveis de interesse: despesas totais (y), avarias de equipamento (x_1), condições metereológicas (x_2) e número de trabalhadores (x_3).

Considerou-se um modelo de regressão múltipla para descrever y em função das restantes variáveis:

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-22,950	7,762		-2,957	,005
	avarias de equipamento	,875	,280	,706	2,409	,021
	condições metereológicas	-,201	,125	-,360	-1,608	,117
	número de trabalhadores	28,277	9,665	,626	2,926	,006

a. Dependent Variable: despesas totais

(a) Diga para que níveis de significância é que a ordenada na origem se pode considerar inferior a zero.

(b) Teste, ao nível de significância 0.05, se o coeficiente associado às avarias de equipamento se pode considerar diferente de 0.5.

(c) Obtenha o *p-value* do teste ao coeficiente das condições metereológicas, em falta na tabela.

(d) Considere os seguintes gráficos e tabelas.

Model Summary^a

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,975 ^a	,951	,947	1,13112

a. Predictors: (Constant), número de trabalhadores, condições metereológicas, avarias de equipamento

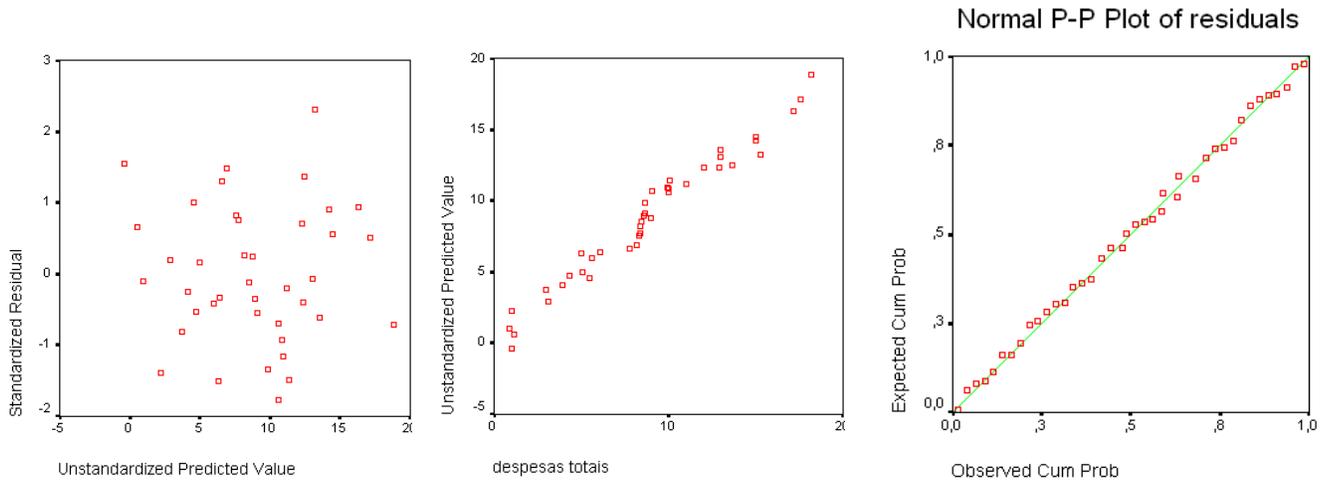
b. Dependent Variable: despesas totais

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	834,301	3	278,100	351,989	,000 ^a
	Residual	28,443	36	,790		
	Total	862,744	39			

a. Predictors: (Constant), número de trabalhadores, condições metereológicas, avarias de equipamento

b. Dependent Variable: despesas totais



- A estimativa do desvio padrão dos erros é $\hat{\sigma} =$
- Com base nos gráficos e/ou tabelas que achar conveniente o que pode afirmar sobre o significado da regressão e a qualidade do ajuste?