



Exame de Recurso de Métodos Estatísticos
Departamento de Matemática – Universidade de Aveiro

Data: 08/07/2004

Duração: 3 horas

Nome: _____ N.º: _____

Curso: _____ Regime: _____

Declaro que desisto _____

Classificação: _____

As cotações deste exame encontram-se na seguinte tabela:

Questão	1			2			3	4		5		6			7				8			
	a	b	c	a	b ₁	b ₂		a	b	a	b	a	b	c	a	b	c ₁	c ₂	a	b	c	d
Cotação	5	5	10	5	5	10	10	10	15	10	15	10	10	10	10	10	10	10	5	5	10	10
Classif.																						

Responda às questões utilizando o espaço reservado para o efeito. Tenha em atenção a clareza e apresentação das suas respostas.

Reuniram-se os deuses que, já desde a época de Viriato, nos têm dado uma mãozinha, para discutir o estado da nação (portuguesa, claro!)... que não deixa de ser preocupante. O que fazer com toda esta gente perdida e desorientada? O que era feito dos tempos áureos das Descobertas que deram a conhecer Portugal ao mundo inteiro? Decidiram-se então por espalhar oráculos aleatoriamente pelo território português, pessoas que serviriam de porta-vozes dos deuses, para aconselhar quem os consultasse.

1. O oráculo de Aveiro é procurado em média por 25 pessoas por dia.

(a) O número de pessoas que o oráculo de Aveiro recebe por dia pode ser bem modelado por uma distribuição _____ de parâmetro(s) _____.

(b) Qual a probabilidade do oráculo receber hoje exactamente 3 pessoas?

(c) Dê um valor aproximado para a probabilidade do oráculo receber entre 30 e 40 pessoas (use o Teorema do Limite Central).

2. O tempo de resposta do oráculo a cada pergunta de um comum mortal é uma variável aleatória com distribuição exponencial.

(a) Determine, pelo método dos momentos, um estimador para o parâmetro desta distribuição e uma sua estimativa, sabendo que o tempo de resposta a 20 perguntas totalizou 200 minutos.

(b) Com base na alínea anterior (se não respondeu, faça $\lambda = 0.1$) e sabendo que a função de distribuição exponencial é dada por $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$ determine:

(b1) a probabilidade de um desgraçado português ter de esperar mais de 10 minutos por uma resposta a uma sua pergunta;

(b2) a probabilidade de um desgraçado português ter de esperar pelo menos mais 10 minutos por uma resposta quando já esperou 10 minutos, $P(X > 20 | X > 10)$.

3. Os deuses pediram ao nosso oráculo um relatório do seu serviço para terem uma ideia mais correcta da utilidade do seu trabalho. O oráculo, com algum receio pelo seu trabalho, verificou que cerca de 27% das 583 pessoas que o procuraram no mês de Maio seguiram o conselho dado pelos deuses através dele. Os deuses já ficavam contentes se 30% das pessoas seguissem os seus conselhos...

Será que os deuses têm razões para ficarem contentes de acordo com o relatório do oráculo? Dê a sua resposta a um nível de significância de 5% e utilizando um dos testes que se seguem (indique qual).

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
resposta	583	,2676	,44308	,01835

One-Sample Test

	Test Value = 0.3					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
	Lower	Upper				
resposta	-1,767	582	,078	-,03242	-,0685	,0036

One-Sample Test

	Test Value = 0.27					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
	Lower	Upper				
resposta	-,132	582	,895	-,00242	-,0385	,0336

Binomial Test

		Category	N	Observed Prop.	Test Prop.	Asymp. Sig. (1-tailed)
resposta	Group 1	seguiu o conselho	156	,27	,3	,047(a,b)
	Group 2	não seguiu o conselho	427	,73		
	Total		583	1,0		

a Alternative hypothesis states that the proportion of cases in the first group < ,3.

b Based on Z Approximation.

One-Sample Kolmogorov-Smirnov Test

		resposta
N		583
Normal Parameters(a,b)	Mean	,2676
	Std. Deviation	,44308
Most Extreme Differences	Absolute	,459
	Positive	,459
	Negative	-,273
Kolmogorov-Smirnov Z		11,094
Asymp. Sig. (2-tailed)		,000

a Test distribution is Normal.

b Calculated from data.

4. O Rocha, tendo ficado um *expert* nos Métodos Estatísticos, graças à *Grande Maratona*, teve emprego garantido com os deuses. Assim, foi por eles contactado para os ajudar nalgumas decisões. Relativamente à capacidade dos portugueses fazerem face a despesas de saúde, o Rocha contactou o Hospital de Aveiro e seleccionou aleatoriamente 25 contas por pagar que totalizavam 2974 euros.

(a) Assumindo que o montante de cada dívida tem distribuição normal e que a experiência do hospital em anos anteriores permite considerar um desvio-padrão populacional de 50 euros, construa um intervalo de confiança a 95% para o montante médio de cada conta não saldada.

- (b) Nas mesmas condições da alínea anterior, poderá afirmar, ao nível de significância de 5%, que o valor médio de cada conta não saldada é superior a 100 euros? Efectue o teste através da região crítica.

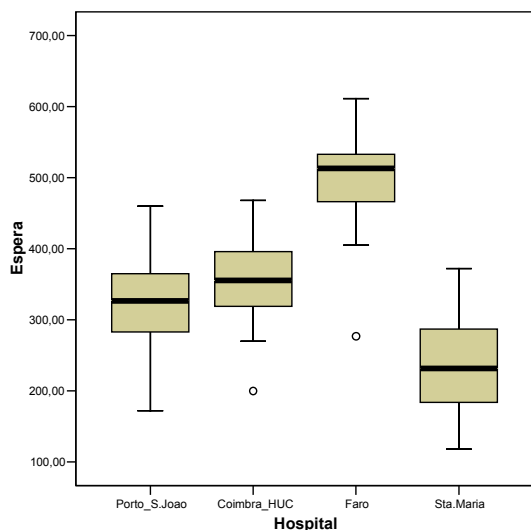
5. Mas o Rocha não ficou por aqui: quis impressionar os deuses com um trabalho mais completo. Quem sabe se poderia ganhar mais... *favores*. Decidiu ir a Coimbra e comparar as famosas listas de espera dos doentes em Coimbra e em Aveiro. Para isso, recolheu em Coimbra dados sobre 40 doentes: tempo de espera até à consulta (X_i^{HC}) e respectivo tempo de espera entre a consulta e a cirurgia (Y_i^{HC}). Mas em Aveiro a coisa foi mais complicada. A grande desorganização levou a que só conseguisse os tempos de espera até à consulta de 40 doentes (X_i^{HA}) e tempos de espera entre a consulta e a cirurgia (Y_i^{HA}) de outros 30 doentes.

Admita que todas as amostras são aleatórias e provenientes de populações normais e considere os seguintes testes de comparação de 2 grupos:

	$\mu_X^{HC} - \mu_Y^{HC}$	$\mu_X^{HA} - \mu_Y^{HA}$
teste- <i>t</i> para amostras emparelhadas		
teste- <i>t</i> para amostras independentes		
teste de Wilcoxon		
teste dos sinais		
teste de Mann-Whitney U		

- (a) Assinale com uma cruz (X) qual o teste que escolheria para comparar:
- o tempo de espera para a consulta e o tempo de espera entre a consulta e a cirurgia no Hospital de Coimbra (HC).
 - o tempo de espera para a consulta e o tempo de espera entre a consulta e a cirurgia no Hospital de Aveiro (HA).
- (b) Considere a variável $\bar{X}^{HA} + \bar{Y}^{HA}$ que estima o tempo médio $\mu_X^{HA} + \mu_Y^{HA}$ que cada doente tem de esperar desde que marca a consulta até ser operado. Diga qual é a sua distribuição não esquecendo de especificar os parâmetros o mais possível.

6. Assustados, quiseram os deuses averiguar se alguns hospitais são mais eficientes do que outros... mais trabalho para o Rocha que percebe mesmo disto. Assim, comparou o tempo de espera por uma cirurgia em vários hospitais do país através da seguinte ANOVA:



ANOVA

Espera

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1090634,467	3	363544,822	91,391	,000
Within Groups	461434,733	116	3977,886		
Total	1552069,200	119			

Multiple Comparisons

Dependent Variable: Espera

	(I) Hospital	(J) Hospital	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Bonferroni	Porto_S.Joao	Coimbra_HUC	-32,96667	16,28473	,271	-76,6791	10,7457
		Faro	-174,33333(*)	16,28473	,000	-218,0457	-130,6209
		Sta.Maria	90,90000(*)	16,28473	,000	47,1876	134,6124
	Coimbra_HUC	Porto_S.Joao	32,96667	16,28473	,271	-10,7457	76,6791
		Faro	-141,36667(*)	16,28473	,000	-185,0791	-97,6543
		Sta.Maria	123,86667(*)	16,28473	,000	80,1543	167,5791
	Faro	Porto_S.Joao	174,33333(*)	16,28473	,000	130,6209	218,0457
		Coimbra_HUC	141,36667(*)	16,28473	,000	97,6543	185,0791
		Sta.Maria	265,23333(*)	16,28473	,000	221,5209	308,9457
	Sta.Maria	Porto_S.Joao	-90,90000(*)	16,28473	,000	-134,6124	-47,1876
		Coimbra_HUC	-123,86667(*)	16,28473	,000	-167,5791	-80,1543
		Faro	-265,23333(*)	16,28473	,000	-308,9457	-221,5209

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

(a) Quantos hospitais foram analisados e qual o tamanho de cada grupo?

(b) Que conclusão tira observando o quadro da ANOVA?

(c) Que conclusões pode tirar a partir da análise de comparações múltiplas pelo método de Bonferroni? Compare essas conclusões com o que se observa nas caixas de bigodes paralelas.

7. Por outro lado, as situações de injustiça social geradas pela recente vaga de imigração de pessoas à procura de melhores condições de vida, colocaram os deuses em sobressalto. As notícias não abonavam em favor dos portugueses... Mas os deuses, antes de fazerem juízos precipitados, recorreram mais uma vez ao Rocha.

Solicitaram-lhe que registasse os salários mensais de trabalhadores da construção civil... mas ninguém queria (ou sabia) falar com ele. Lá conseguiram reunir informação de 8 portugueses (amostra A) e 6 imigrantes de leste (amostra B).

(a) Relativamente ao quadro seguinte que resulta de aplicar o teste de Kolmogorov-Smirnov à amostra B, indique as hipóteses que estão a ser testadas e diga se rejeita a hipótese nula, ao nível de significância de 0.05.

One-Sample Kolmogorov-Smirnov Test(c)

		salario
N		6
Normal Parameters(a,b)	Mean	285,8333
	Std. Deviation	54,07556
Most Extreme Differences	Absolute	,282
	Positive	,282
	Negative	-,198
Kolmogorov-Smirnov Z		,690
Asymp. Sig. (2-tailed)		,727
Exact Sig. (2-tailed)		,985
Point Probability		,000

a Test distribution is Normal b Calculated from data c nacionalidade = B

(b) Assinale a frase que melhor se adequa a comentar a *performance* do teste da alínea anterior.

Dada a reduzida dimensão da amostra, o teste de KS tem muito baixa potência e portanto tenderá a aceitar a hipótese de Normalidade mesmo sem ela ser verdadeira.

Dado o elevado valor do *p-value*, não há dúvida que os dados provêm de uma população Normal.

(c) Pretende-se averiguar se os portugueses são mais bem pagos do que os imigrantes. Considere os seguintes testes efectuados pelo SPSS:

Mann-Whitney

salario	B	N	6
		Mean Rank	5,08
		Sum of Ranks	30,50
	A	N	8
		Mean Rank	9,31
		Sum of Ranks	74,50
Total		N	14

Test Statistics^b

	salario
Mann-Whitney U	9,500
Wilcoxon W	30,500
Z	-1,874
Asymp. Sig. (2-tailed)	,061
Exact Sig. [2*(1-tailed Sig.)]	,059 ^a
Exact Sig. (2-tailed)	,064
Exact Sig. (1-tailed)	,032
Point Probability	,005

a. Not corrected for ties.

b. Grouping Variable: nacionalidade

t-test

Paired Samples Test

		Paired Differences				t	df	Sig. (2-tailed)	
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	salario - nacionalidade	338,78571	104,54015	27,93953	278,42603	399,14540	12,126	13	,000

Wilcoxon Signed Rank test

Ranks

nacionalidade - salario	Negative Ranks	N	14 ^a
		Mean Rank	7,50
		Sum of Ranks	105,00
	Positive Ranks	N	0 ^b
		Mean Rank	,00
		Sum of Ranks	,00
	Ties	N	0 ^c
Total		N	14

a. nacionalidade < salario

b. nacionalidade > salario

c. nacionalidade = salario

Test Statistics^b

	nacionalidade - salario
Z	-3,296 ^a
Asymp. Sig. (2-tailed)	,001
Exact Sig. (2-tailed)	,000
Exact Sig. (1-tailed)	,000
Point Probability	,000

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

Sign test

Frequencies

		N
nacionalidade - salario	Negative Differences ^a	14
	Positive Differences ^b	0
	Ties ^c	0
	Total	14

a. nacionalidade < salario

b. nacionalidade > salario

c. nacionalidade = salario

Test Statistics^b

	nacionalidade - salario
Exact Sig. (2-tailed)	,000 ^a
Exact Sig. (1-tailed)	,000
Point Probability	,000

a. Binomial distribution used.

b. Sign Test

(c₁) Diga, justificando, qual dos testes escolheria para dar resposta ao pretendido.

(c₂) Diga, ao nível de significância de 0.1, a que conclusão chega, indicando o *p-value* do teste e comparando-o com o nível de significância.

8. Desesperados com o panorama nacional, os deuses voltaram-se para a cultura com a esperança de verem algo de menos negativo. Investigaram as contas bancárias de 14 agregados familiares para conhecerem os seus rendimentos mensais *per capita* (X) e saberem quanto gastavam em bens e serviços culturais (Y) (a tarefa não foi dada ao Rocha por causa do sigilo bancário...).

Realizada uma análise de regressão linear no SPSS, obtiveram, entre outros, os seguintes *outputs*

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,836(a)	,699	,674	58,28862

a Predictors: (Constant), Rendimento

Coefficients(a)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	30,735	26,735		1,150	,273
	Rendimento	,100	,019	,836		

a Dependent Variable: GastosCultura

(a) Escreva a equação da recta estimada.

(b) Para que níveis de significância é que a ordenada na origem se pode considerar significativamente diferente de zero?

(c) Preencha os valores que faltam na tabela.

(d) Qual a proporção de variabilidade de Y explicada por X?

Pois é, se não forem os deuses a ajudar... estamos mal.
E assim vai Portugal!