



## Exame Final de Métodos Estatísticos

Data: 09/06/2003

Duração: 2h00

Número: \_\_\_\_\_ Nome: \_\_\_\_\_

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

Número de folhas suplementares: \_\_\_\_\_ Declaro que desisto: \_\_\_\_\_

Esta prova consiste em 5 questões de resposta aberta. As cotações encontram-se na margem esquerda do enunciado.

Cada folha suplementar entregue desconta 0.2 na cotação total.

As aulas terminaram e o fim do ano lectivo aproxima-se. Na Universidade Fictícia existem 36 alunos finalistas do curso de Belas Artes. A Comissão de Curso está a organizar uma viagem de finalistas, estando já decidido que se fará um *Interrail*, forma fácil e económica de combinar diversão e convívio com visitas às mais interessantes obras de arte.

Os países abrangidos encontram-se classificados em oito zonas (A, B, C, D, E, F, G e H) e o *Interrail* pode ser adquirido para 1, 2 ou 3 zonas.

1. A Comissão de Curso considera que devem ser adquiridos bilhetes de duas zonas, mas ainda não decidiu quais delas serão incluídas na viagem. Para tomar tal decisão, pede a cada um dos 36 finalistas que indique a zona que mais lhe interessa.

Pela riqueza artística e cultural de alguns dos países que compõem a zona G (Itália, Eslováquia, Grécia e Turquia) os alunos têm-se mostrado muito interessados em visitá-la. Com base nas opções dos finalistas de anos anteriores considera-se que a probabilidade  $p$  de um aluno escolher a zona G é igual a 0.6.

- (a) Identifique a distribuição da variável aleatória  $X$  que representa o número de alunos finalistas que escolhem a zona G, de entre os 36 que decidiram fazer a viagem.

(1.0)

|

- (b) Calcule a probabilidade de todos optarem pela referida zona.

(1.0)

|

- (c) Uma zona é automaticamente seleccionada como destino da viagem, se for escolhida por pelo menos 18 alunos finalistas. Utilizando o Teorema do Limite Central, calcule a probabilidade aproximada da zona G ser automaticamente seleccionada.

(2.0)

|

- (d) Pretende-se saber se a proporção de alunos que escolhe a zona G se mantém igual ao passado, ou seja, se se pode continuar a considerar como sendo 0.6. Sabendo que neste grupo de 36 finalistas, 18 escolheram a zona G, obtenha uma estimativa para  $p$  utilizando o método dos momentos.

(1.5)

|

- (e) Forneça um intervalo de confiança a 95% confiança para  $p$ . (Se não resolveu a alínea anterior considere  $\hat{p} = 0.5$ .) Com base nesse intervalo teste, ao nível de significância 5%, as hipóteses:

$$H_0 : p = 0.6 \quad vs \quad H_1 : p \neq 0.6.$$

(2.0)

|

Após muita discussão e ponderação, a Comissão de Curso decidiu que a viagem de finalistas teria como destino a zona F (Portugal, Espanha e Marrocos) e a zona G (Itália, Eslovénia, Grécia e Turquia).

2. A cidade espanhola de Barcelona possui uma das maiores mostras de arquitectura modernista do Mundo salientando-se, por exemplo, a obra de Gaudí e a famosa Boqueria (mercado de rua, nas Ramblas). Nas visitas a alguns monumentos, os alunos tiraram fotografias e registaram níveis de apreciação. Os níveis variam de 1 a 10, sendo 10 o nível de maior apreciação.

Em relação à Sagrada Família de Gaudí e à Catedral de Barcelona pretende-se averiguar se o grau de apreciação dos alunos é significativamente diferente para os dois monumentos. Considere que  $\mu_A$  e  $\mu_B$  representam a média (ou mediana) dos níveis de apreciação destes dois monumentos, respectivamente.

Efectuaram-se os seguintes testes bilaterais e registaram-se os p-values associados:

Teste	p-value (bilateral)
teste t de comparação de médias (amostras independentes)	0.002
teste t de comparação de médias (amostras emparelhadas)	0.004
Mann Whitney – U	0.025
Wilcoxon Signed Ranks	0.031

- (a) Indique, justificando, o teste que escolheria.

(1.0)

|

(1.5)

- (b) Escreva as hipóteses do teste e conclua.

|

3. O grupo de finalistas ficou três dias em Liubliana, capital da Eslovénia. A cerca de 50 km da capital fica o lago Bled, uma autêntica pérola nos Alpes eslovenos. Rodeado de montanhas, a côr e a transparência da água não têm explicação, sendo indispensável um passeio de “pletna” (embarcação típica) até à ilha existente no meio do lago, com a sua igreja ao cimo de uma escadaria.

Após o passeio de pletna, todos fizeram questão de subir os 99 degraus que dão acesso à igreja. Os mais desportistas do grupo transformam essa subida numa corrida. Os 15 tempos registados, em segundos, foram

18 20 18 20 19 16 20 13 16 30 29 15 22 12 20

cujas características amostrais mais relevantes são:

mínimo	média	mediana	desvio padrão corrigido	máximo
12	19.2	19	5.03	30

Considere que a variável  $X$ =“tempo que demora a subida da escadaria (em seg)” é bem modelada por uma distribuição normal.

- (a) Teste, ao nível de significância de 0.05, se o tempo médio de subida é significativamente diferente de 18. Ou seja, teste as hipóteses

$$H_0 : \mu = 18 \quad vs \quad H_1 : \mu \neq 18$$

(2.0)

- (b) Determine o *p-value* do teste anterior.

(1.5)

4. Como se sabe, em Portugal, as compras a marroquinos não são feitas sem regatear os preços... Em Marrocos, muito menos! No último dia em Marrocos, os 36 finalistas foram comprar recordações. Com o espírito de competição que os caracteriza dividiram-se aleatoriamente em três grupos (homogêneos) e percorreram a feira em busca dos melhores descontos. Cada finalista registou o desconto conseguido na sua primeira compra. No sentido de avaliar possíveis diferenças entre os descontos médios dos grupo, efectuou-se uma ANOVA com base nesses valores. Complete a tabela de ANOVA que foi obtida

(2.0)

**ANOVA**

DESCON

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1,436	.....	.....	.....	,417
Within Groups	26,409	.....	.....		
Total	.....	.....			

5. Na viagem de regresso a Lisboa decidiram contabilizar o *total de gastos*. Suspeita-se que os gastos de cada um dependem de vários factores tais como: a remuneração média que cada aluno tem por mês (seja na forma de mesada, bolsa ou vencimento); a poupança que cada um fez durante o último ano; e o número de viagens ao estrangeiro que cada um já efectuou no passado. Assim, cada aluno contabilizou os valores referentes as estas variáveis aos quais se ajustou um modelo de regressão linear múltipla. Considere  $x_1$  =remuneração,  $x_2$  =poupança e  $x_3$  =número de viagens. As seguinte tabela contém alguns resultados dessa análise de regressão.

**Coefficients<sup>a</sup>**

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3,259	84,178		,039	,969
	remuneração	1,206	,198	,741	6,104	,000
	poupança	8,425E-02	,054	,191	1,574	,125
	nº de viagens	-17,013	9,814	-,204	-1,733	,093

a. Dependent Variable: gasto total

- (a) Escreva a equação da superfície de regressão estimada, ou seja, na equação  $y = \hat{\beta}_0 + \hat{\beta}_1x_1 + \hat{\beta}_2x_2 + \hat{\beta}_3x_3$  substitua os estimadores pelas suas estimativas.

(1.0)

- (b) Diga se existe algum coeficiente de regressão que possa ser considerado significativamente **inferior** a zero (teste **unilateral**) não esquecendo de indicar para que níveis de significância é que tal se observa.

(1.5)

- (c) Como sabe uma forma de avaliar o significado da regressão consiste em comparar o valor da estimativa do desvio padrão dos gastos totais sem tomar em consideração as restantes variáveis com a estimativa do desvio padrão dos erros do modelo de regressão. Com base na tabela de ANOVA da regressão

**ANOVA<sup>b</sup>**

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	115790,37	3	38596,791	13,567	,000 <sup>a</sup>
	Residual	91038,135	32	2844,942		
	Total	206828,51	35			

a. Predictors: (Constant), nº de viagens, remuneração, poupança

b. Dependent Variable: gasto total

forneça estas estimativas para os dados em causa.

(2.0)