



UNIVERSIDADE DE COIMBRA
FACULDADE DE CIÊNCIAS E TECNOLOGIA
Departamento de Engenharia Informática
Pólo II - Pinhal de Marrocos, 3030 Coimbra - Portugal

Exame de Métodos Experimentais em Informática (Recurso)

27 de Janeiro de 2016

Tempo máximo: 1h:45m

Nota: Os estudantes podem utilizar o seu material em papel para consulta durante o exame. O único dispositivo eletrónico permitido é a calculadora, desde que não esteja ligada à Internet.

Ponderações: 1 a) 15% 1 b) 15% 1 c) 10% 1 d) 10% 2 a) 12.5% 2 b) 12.5% 2 c) 12.5% 2 d) 12.5%

1. a) Um grupo de estudantes de Engenharia Informática quer testar se a temperatura a partir da qual o CPU de uma determinada marca (marca A) começa a apresentar problemas é inferior a 110°C. Eles efetuaram testes em 20 portáteis da marca A e constataram que, em média, o equipamento não suporta uma temperatura superior a 109°C com um desvio padrão de 3°C. Assumindo que a distribuição da temperatura máxima suportada pelo CPU dessa marca segue uma distribuição próxima da normal, o que se pode concluir a partir dos resultados da experiência para um nível de significância de 5%? Apresente os quatro passos do teste de hipóteses e comente o tipo de erro associado à decisão.

Nota: Considere a seguinte informação, em que a probabilidade corresponde à área debaixo da curva da respetiva distribuição de $-\infty$ ao valor indicado na tabela.

Probabilidade	Valor da distribuição normal estandardizada	Valor da distribuição t de Student para 19 graus de liberdade	Valor da distribuição t de Student para 28 graus de liberdade
2.5%	-1.959964	-2.306004	-2.048407
5.0%	-1.644854	-1.859548	-1.701131
47.5%	-0.062707	-0.064701	-0.063271
52.5%	0.062707	0.064701	0.063271
95.0%	1.644854	1.859548	1.701131
97.5%	1.959964	2.306004	2.048407

b) O mesmo grupo de estudantes leu uma notícia no Facebook que o CPU de uma determinada marca (marca B) consegue suportar uma temperatura superior ao da marca A. Para verificarem se é mesmo verdade, o grupo de estudantes adquiriu 10 portáteis da marca B e efetuou vários testes, tendo obtido uma média de 112°C e desvio padrão 2°C. Assumindo que a distribuição da temperatura máxima suportada pelo CPU da marca B segue igualmente uma distribuição próxima da normal, e dado os resultados obtidos com a experiência na marca A descrita na alínea anterior, o que se pode concluir para um nível de significância de 5%? Apresente os quatro passos do teste de hipóteses e comente o tipo de erro associado à decisão.

c) Se os pressupostos descritos na alíneas anteriores não forem válidos (por exemplo, se a distribuição não for normal), uma possibilidade seria recorrer a um teste de aleatorização (*randomization test*). Descreva como este teste poderia ser efetuado para responder à questão anterior.

d) As duas firmas (da marca A e B), ao saberem da experiência efetuada pelos estudantes, disponibilizaram 40 portáteis de cada marca para repetirem a experiência. Os estudantes mudaram as condições experimentais e contaram quantos CPUs da marca A e B é que suportavam um valor superior ou igual a 110°C. Obtiveram os seguintes resultados: 20 da marca A e 22 da marca B. Calcule o intervalo de confiança para a diferença entre estes dois resultados, assumindo um nível de confiança de 95%. Explique o significado deste intervalo e relacione com o resultado obtido na alínea b).

2. **a)** Um programador está a desenvolver um algoritmo para resolver um problema de caminho mais curto que surge numa aplicação de navegação por telemóvel. Embora seja um especialista em três linguagens de programação (C, C++ e Java), não sabe qual destas deve ser utilizada para a aplicação. Deste modo, decidiu implementar o algoritmo nas três linguagens, registou os tempos obtidos em 5 telemóveis diferentes e obteve os seguintes resultados (em segundos):

C	C++	Java
58	51	77
45	68	72
40	64	78
41	63	73
41	62	75

Com base nesta informação, preencha a seguinte tabela ANOVA (apresente os cálculos):

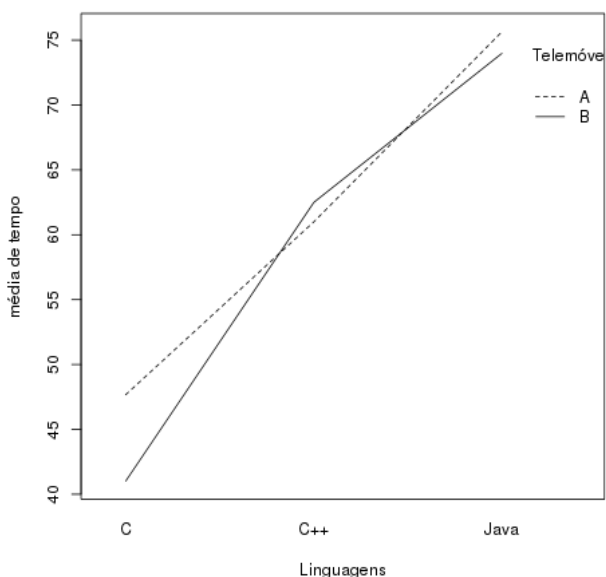
	SS	df	MS	F	P-value
Between					1.37e-05
Within					
Total					

- b)** A que conclusão chega com a tabela ANOVA para um valor de significância de 1%?

- c)** O teste emparelhado de t de Student (*pairwise t-test*) com correção para múltiplas comparações através do método de Bonferroni reporta valores de p (*p-values*) que estão na seguinte tabela. A que conclusão chega para um nível de significância de 1%? Porque necessita de um método de correção para múltiplas comparações? O que poderia acontecer se esta correção não fosse utilizada?

	C	C++
C++	0.0023	-
Java	1e-05	0.0107

- d)** Ao rever as condições experimentais descritas nas alíneas anteriores, o programador verificou que tinha duas marcas diferentes de telemóveis. Por essa razão, repetiu a análise e obteve o seguinte gráfico de interação (*interaction plot*) e a seguinte tabela ANOVA. Dadas as respostas às alíneas anteriores, a que conclusão deverá chegar o programador para um nível de significância de 1% com esta nova informação?



	SS	df	MS	F	P-value
Linguagem	2258.5	2	1129.3	28.724	0.000124
Telemóvel	18.7	1	18.7	0.475	0.508037
Interaction	40.7	2	20.3	0.517	0.612735
Within	353.8	9	39.3		
Total	2671.7	14			

Errata

Os valores da distribuição t de Student para 19 graus de liberdade na questão 1a) são:

2.5% : -2.093024

5% : -1.729133

47.5%: -0.06354029

52.5%: 0.06354029

95% : 1.729133

97.5% : 2.093024