



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 (Normal)

7 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.

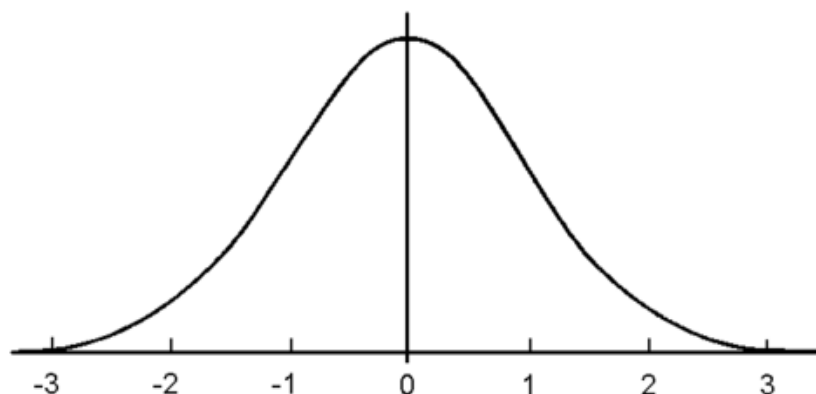
Ponderações: 1 a) 20% 1 b) 5% 2 a) 15% 2 b) 15% 2 c) 5% 2 d) 20% 3 a) 10% 3 b) 10%

1a) Um novo sistema de base de dados vai ser implementado num banco com o objetivo de aumentar a rapidez das consultas feitas pelos clientes através da Internet. Existem vários tipos de consultas que são importantes assegurar da forma mais rápida possível: i) consulta de saldo da conta de depósito à ordem à escolha do cliente, ii) consulta de movimentos no período escolhido pelo cliente, e iii) consulta da posição global, que calcula o saldo de todas as contas de depósito à ordem, prazo, poupanças e investimentos que possui. O tempo de cada tipo de consulta pode depender do número de movimentos e de contas e do intervalo de tempo considerado. Com o objetivo de testar a nova base de dados, o departamento de sistemas de informação do Banco criou várias contas e movimentos em ambos os sistemas (atual e novo) e efetuou uma experiência para testar a rapidez do novo sistema relativamente ao sistema atual. Identifique os vários elementos para o desenho da experiência (definição do problema, variáveis e hipóteses), justificando as suas escolhas.

b) Uma das consultas que foi testada durante a experiência demorou muito mais tempo do que as restantes consultas com o mesmo número de movimentos. Quais as alternativas disponíveis para lidar com este resultado?

2a) Assuma que adquiriu um novo sistema de navegação cujo o tempo para encontrar um sinal válido de GPS segue uma distribuição desconhecida com média de 10 segundos e desvio padrão de 6 segundos. Se recolher 36 medições em locais e períodos de tempo escolhidos aleatoriamente, qual a probabilidade que o tempo médio destas medições demore entre 11 e 12 segundos? Justifique as suas escolhas.

Nota: Não necessita de calcular explicitamente a probabilidade, indique unicamente os cálculos a efetuar e ilustre o valor da probabilidade no seguinte gráfico da distribuição normal standardizada (transcreva o gráfico na folha de teste)



b) Para validar a informação disponibilizada pelo fabricante do sistema de navegação, efetuou 9 medições para encontrar um sinal válido de GPS em locais e períodos de tempo escolhidos aleatoriamente, tendo obtido uma média de 10.5 segundos e um desvio padrão de 5 segundos. Qual o intervalo de confiança de 95% para o tempo médio de encontrar o sinal válido de GPS? Explique o significado deste intervalo.

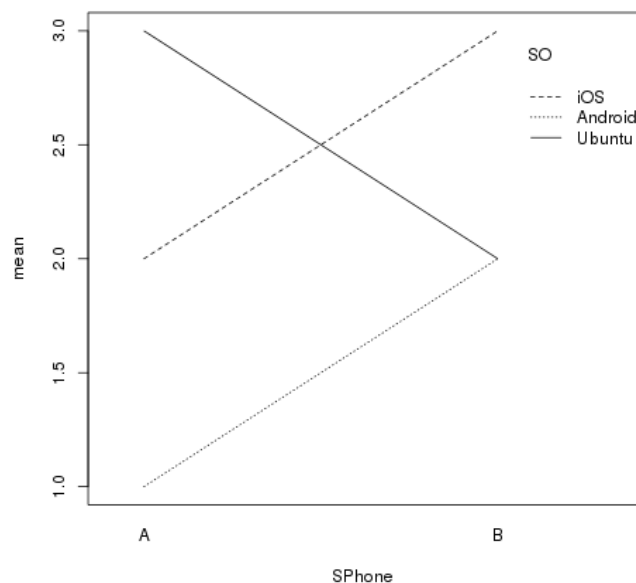
Nota: Considere a seguinte informação para apoiar nos cálculos, em que a probabilidade corresponde à área debaixo da curva da respetiva distribuição de $-\infty$ ao valor indicado na tabela.

Probabilidade	Valor da distribuição normal estandardizada	Valor da distribuição t de Student para 8 graus de liberdade
2.5%	-1.959964	-2.306004
5.0%	-1.644854	-1.859548
47.5%	-0.06270678	-0.0647005
52.5%	0.06270678	0.0647005
95.0%	1.644854	1.859548
97.5%	1.959964	2.306004

c) Seria possível obter um intervalo de menor amplitude para o mesmo nível de confiança? Justique a resposta.

d) Assumindo um nível de significância de 5%, deve ou não confiar nos resultados do fabricante? Apresente os quatro passos do teste de hipóteses e utilize a tabela da alínea anterior para os cálculos. Comente o tipo de erro associado à sua decisão.

3.a) Uma companhia desenvolveu um novo algoritmo para procurar palavras em ficheiros de *smartphones* e pretende testar a sua rapidez em três sistemas operativos: Android, iOS e Ubuntu Touch e em dois tipos de dispositivos (A e B). Foram efetuadas várias medições e obteve-se o seguinte gráfico de interação (*interaction plot*) dos tempos obtidos. Que conclusões é possível obter do gráfico?



b) A seguinte tabela apresenta os resultados do ANOVA de acordo com um modelo de dois fatores e interações. Que conclusões retira da tabela apresentada para um valor de significância de 5%? Compare com as conclusões obtidas na alínea anterior. Se os pressupostos deste teste estiverem corretos, qual deverá ser o próximo passo da análise?

	Df	Sum Sq	Mean Sq	F value	P-value
Factor SO	2	23.333	11.6667	17.5	1.384e-06
Factor SPhone	1	1.667	1.667	2.5	0.1197
Interação	2	23.333	11.6667	17.5	1.384e-06
Resíduos	54	36.000	0.6667		