

## Grupo A

### QUESTÃO 1+++++

(1Ch 2007-06-19)

1. Mostre que se  $\hat{\theta}$  é um estimador centrado do parâmetro  $\theta$  e  $\text{var}[\hat{\theta}] > 0$ , então  $\hat{\theta}^2$  não é um estimador centrado de  $\theta^2$ .

(1Ch 2006-06-19)

1. Mostre a seguinte igualdade

$$EQM = E[(T - \mu)^2] = \text{Var}[T] + [E(T) - \mu]^2$$

para o Erro Quadrático Médio  $EQM$ .

(2Ch 2006-07-04)

1. Na Análise de Variância de uma entrada, para amostras de tamanho diferente, a partir da relação

$$Y_{ij} - \bar{Y} = (Y_{ij} - \bar{Y}_i) + (\bar{Y}_i - \bar{Y}) \text{ mostre que:}$$

$$\sum \sum (Y_{ij} - \bar{Y})^2 = \sum \sum (Y_{ij} - \bar{Y}_i)^2 + \sum_i n_i (\bar{Y}_i - \bar{Y})^2 \quad .$$

(RCh 2006-07-24)

1. Dado um conjunto de pontos  $(X_i, Y_i)$ , considere o seguinte modelo de regressão linear

$$Y_i = \alpha + \beta(X_i - \bar{X}) + \varepsilon_i$$

- a) Mostre que a soma dos resíduos é igual a zero;
- b) Tomando como variável dependente os resíduos obtidos  $\varepsilon_i$  e variável independente  $X_i$ , mostre que a recta de regressão de mínimos quadrados tem declive nulo.

(1Ch 2005-06-23)

1. Considerando a seguinte igualdade,

$$(y_{ij} - \bar{y}_{..}) = (\bar{y}_{i.} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.})$$

mostre que no planeamento completamente aleatório

$$STQ = SQT + SQR$$

isto é, que a Soma Total dos Quadrados é igual à Soma dos Quadrados dos Tratamentos mais a Soma dos Quadrados dos Resíduos,

$$\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = n \sum_{i=1}^k (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2$$

(2Ch 2005-07-06)

1. Dado um conjunto de pontos  $(X_i, Y_i)$ , considere o seguinte modelo de regressão:

$$Y_i = \beta_0 + \beta_1 X_i^2 + \varepsilon_i$$

Determine as estimativas de  $\beta_0$  e  $\beta_1$  pelo método dos mínimos quadrados.

(RCh 2005-07-26)

1. Dado um conjunto de pontos  $(X_i, Y_i)$ , considere o seguinte modelo de regressão linear

$$Y_i = \alpha + \beta(X_i - \bar{X}) + \varepsilon_i$$

- a) Mostre que a soma dos resíduos é igual a zero;  
b) Tomando como variável dependente os resíduos obtidos  $\varepsilon_i$  e variável independente  $X_i$ , mostre que a recta de regressão de mínimos quadrados tem declive nulo.

(1Ch 2004-06-26)

1. Mostre que, para um conjunto de valores  $(X_i, Y_i)$ , no modelo de regressão linear simples

$$\hat{Y}_i = \hat{\alpha}' + \hat{\beta}X_i$$

a intercepção na origem  $\hat{\alpha}'$  pode ser estimada como

$$\hat{\alpha}' = \bar{Y} - \hat{\beta}\bar{X} = \sum_i \left( \frac{1}{n} - \frac{\bar{X}(X_i - \bar{X})}{\sum_i (X_i - \bar{X})^2} \right) Y_i.$$

(2Ch 2004-07-09)

1. Mostre que a soma dos resíduos da recta de regressão de mínimos quadrados é igual a zero.

(RCh 2004-07-24)

1. No Planeamento Completamente Aleatório mostre a seguinte igualdade:

$$STQ = SQT + SQR$$

$$\sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = n \sum_{i=1}^k (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^k \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2$$

(1Ch 2003-06-12)

1. Dado um conjunto de pontos  $(X_i, Y_i)$ , considere o seguinte modelo de regressão linear

$$Y_i = \alpha + \beta(X_i - \bar{X}) + \varepsilon_i$$

- a) Mostre que a soma dos resíduos é igual a zero;  
b) Tomando como variável dependente os resíduos obtidos  $\varepsilon_i$  e variável independente  $X_i$ , mostre que a recta de regressão de mínimos quadrados tem declive nulo

(2Ch 2003-06-30)

1. Mostre que, para o caso de duas amostras do mesmo tamanho,

Amostra 1	$y_{11}$	$y_{12}$	...	$y_{1n}$
Amostra 2	$y_{21}$	$y_{22}$	...	$y_{2n}$

a estatística  $F = \text{MQT}/\text{MQR}$ , (MQT - Média do Quadrado dos Tratamentos, MQR - Média do Quadrado dos Resíduos) da análise da variância para comparação entre tratamentos, é igual ao quadrado da estatística T para a comparação entre as médias de duas amostras,  $T = \frac{(\bar{y}_1 - \bar{y}_2)}{Sp \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ .

(1Ch 2002-06-03)

1. Mostre que na regressão linear simples, em que  $Y_i = \alpha + \beta X_i + \varepsilon_i$

- O ponto  $(\bar{X}, \bar{Y})$  está exactamente na linha de regressão de mínimos quadrados;
- A partir da igualdade, para qualquer  $i$ , dada por  $(Y_i - \bar{Y}) = (\hat{Y} - \bar{Y}) + (Y_i - \hat{Y}_i)$  a variação total é dada por  $\sum_i (Y_i - \bar{Y})^2 = \sum_i (\hat{Y} - \bar{Y})^2 + \sum_i (Y_i - \hat{Y}_i)^2$

(2Ch 2002-06-20)

1. No modelo de regressão linear simples, mostre que:

- a soma dos valores observados da variável dependente é igual à soma dos valores estimados,  $\sum_i Y_i = \sum_i \hat{Y}_i$ ;
- a regressão dos erros  $\varepsilon_i$  em função da variável independente  $X_i$  tem declive nulo.

(1Ch 2007-06-19)

2. Suponha que a pressão diastólica num grupo etário de uma população Normal tem média igual a 78mmHg e desvio padrão 9mmHg. Calcule a probabilidade de que em uma amostra de tamanho 16, a pressão diastólica média seja maior que 81mmHg.

(1Ch 2006-06-19)

2.  $X_1$  e  $X_2$  são v.a. independentes contínuas com função densidade assim definida:

$$f_{X_i}(x) = \frac{x^{i-1} e^{-x/\theta}}{\theta^i} \quad x > 0, \theta > 0 \quad i = 1, 2$$

Considere os seguintes estimadores de  $\theta$ :

$$T_1 = \frac{1}{3}(X_1 + X_2) \quad \text{e} \quad T_2 = \frac{1}{2}(X_1 + X_2)$$

Diga qual deles é centrado.

(2Ch 2006-07-04)

2. A distribuição de Pareto, usada para modelar distribuição de renda, tem função de distribuição cumulativa dada por

$$F(x) = \begin{cases} 1 - \left(\frac{\beta}{x}\right)^\alpha, & x \geq \beta \\ 0, & x < \beta \end{cases}$$

onde  $\alpha, \beta > 0$ . Portanto, sua função de densidade é

$$f(x) = \alpha \beta^\alpha x^{-(\alpha+1)}, x \geq \beta.$$

Sejam  $X_1, X_2, \dots, X_n$  i.i.d. com distribuição de Pareto com parâmetro  $\beta$  conhecido e igual a 1. Obtenha o estimador de máxima verosimilhança de  $\alpha$ .

(RCh 2006-07-24)

2. Sejam  $X_1, X_2, \dots, X_7$  os elementos de uma amostra aleatória de uma população com média  $\mu$  e variância  $\sigma^2$ . Considere os seguintes estimadores de  $\mu$ :

$$\hat{\theta}_1 = \frac{X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7}{7} \quad \text{e} \quad \hat{\theta}_2 = \frac{2X_1 - X_6 + X_4}{2}$$

(1Ch 2005-06-23)

2. Seja  $T$  um estimador do parâmetro  $\theta$ . A tendência desse estimador em relação ao parâmetro é dada por  $t_T(\theta) = E[T] - \theta$ . Mostre que

$$MQE = E[(T - \theta)^2] = Var[T] + (t_T(\theta))^2.$$

2. Considere uma amostra aleatória de tamanho  $2n$  retirada de uma população. Sejam  $\bar{X}_1 = \frac{\sum_{i=1}^{2n} X_i}{2n}$  e  $\bar{X}_2 = \frac{\sum_{i=1}^n X_i}{n}$  dois estimadores de  $\mu$ . Qual será o melhor estimador de  $\mu$ ? Justifique.

(RCh 2005-07-26)

2. Sejam  $X_1, X_2, \dots, X_7$  os elementos de uma amostra aleatória de uma população com média  $\mu$  e variância  $\sigma^2$ . Considere os seguintes estimadores de  $\mu$ :

$$\hat{\theta}_1 = \frac{X_1 + X_2 + X_3 + X_4 + X_5 + X_6 + X_7}{7} \quad \text{e} \quad \hat{\theta}_2 = \frac{2X_1 - X_6 + X_4}{2}$$

- a) Serão os estimadores não tendenciosos?  
b) Qual será o estimador preferível? Justifique.

(1Ch 2004-06-26)

2. Seja  $x_1, x_2, \dots, x_n$  uma amostra aleatória de uma distribuição de Poisson, com média  $\lambda$ . Considere os quatro estimadores de  $\lambda$ ,

$$\hat{\lambda}_1 = \bar{x} \quad \hat{\lambda}_2 = n(x_1 + x_2 + \dots + x_n)$$

$$\hat{\lambda}_3 = \frac{x_1 + x_2}{2} \quad \hat{\lambda}_4 = \frac{x_1}{n}$$

Qual ou quais dos estimadores é não enviesado? De entre estes qual preferiria? Justifique.

(2Ch 2004-07-09)

2. Considere a seguinte função,

$$f(x) = \frac{e^{-x}}{1 - e^{-\theta}} \quad 0 < x < \theta$$

Encontre o estimador de máxima verosimilhança de  $\theta$ , considerando uma amostra de tamanho  $n$ .

(RCh 2004-07-24)

2. Seja  $X$  uma variável aleatória com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} (\alpha + 1)x^\alpha & \text{se } 0 < x < 1 \\ 0 & \text{caso contrário} \end{cases}$$

Encontre o estimador de máxima verosimilhança para  $\alpha$ , baseado numa amostra aleatória de dimensão  $n$ .

(1Ch 2003-06-12)

2. Considere a seguinte função de densidade de probabilidade,

$$f(x) = \frac{\lambda}{2} e^{-\lambda|x|} \quad \lambda > 0 \quad -\infty < x < \infty$$

- a) Encontre o estimador de máxima verosimilhança para o parâmetro  $\lambda$ .
- b) Encontre o valor esperado  $E[|X|]$ .
- c) Mostre que o estimador de máxima verosimilhança é não tendencioso.

(2Ch 2003-06-30)

2. Seja  $x_1, \dots, x_{20}$  uma amostra aleatória com dimensão  $n = 20$ , de uma distribuição normal com média  $\mu$  desconhecida e variância conhecida  $\sigma^2 = 5$ . Pretende-se testar

$$H_0 : \mu = 7$$

$$H_1 : \mu > 7$$

- a) Qual a região de rejeição para um nível de significância  $\alpha = 0.05$ ?
- b) Encontre a potência do teste para os seguintes valores da hipótese alternativa,  $\mu = 7.5$ ,  $\mu = 8.0$ ,  $\mu = 8.5$ ,  $\mu = 9.0$ .
- c) Qual o menor valor para a dimensão da amostra tal que com  $\alpha = 0.05$  o teste tenha uma potência de pelo menos 0.8 quando  $\mu = 8$ ?

(1Ch 2002-06-03)

2. Seja  $X_1, X_2, \dots, X_7$  uma amostra aleatória de uma população com média  $\mu$  e variância  $\sigma^2$ .

Considere os seguintes estimadores para  $\mu$ :

$$\theta_1 = \frac{X_1 + X_2 + \dots + X_7}{7} \quad \theta_2 = \frac{2X_1 - X_6 + X_4}{2}$$

- a) Ambos os estimadores são não enviesados?
- b) Qual o melhor estimador? Em que sentido?

(2Ch 2002-06-20)

2. Suponha que tem uma amostra aleatória de dimensão  $2n$  de uma população com média  $\mu$  e variância  $\sigma^2$ . Considere os seguintes estimadores para  $\mu$ :

$$\theta_1 = \frac{1}{2n} \sum_{i=1}^{2n} X_i \quad \theta_2 = \frac{1}{n} \sum_{i=1}^n X_i$$

- a) Qual o melhor estimador? Justifique a sua resposta.

3. A empresa produtora do creme protector solar MORENAÇA deseja testar se uma nova fórmula, recentemente proposta pelo departamento de desenvolvimento, é mais eficaz do que a actualmente adoptada. Para tal seleccionou uma amostra aleatória de 20 pessoas, dividindo-as em 2 grupos de 10. A um grupo aplicou a fórmula actual, e ao outro aplicou a nova fórmula. Os cremes baseados nas duas fórmulas foram aplicados nas costas destes indivíduos que foram submetidos a radiações solares controladas. Os resultados do teste – que exprimem o grau de queimadura solar da pele – apresentam-se na tabela seguinte.

	Fórmula actual	Nova Fórmula
<b>Médias</b>	34,4	30,9
<b>Desvio padrão</b>	11,0	11,3
<b><i>n</i></b>	10	10

Teste ao nível de significância de 5%, se o creme baseado na nova fórmula é um protector mais eficaz do que o actual. Quais os pressupostos utilizados?

(1Ch 2006-06-19)

3. Um determinado fornecedor de acesso à Internet tem um pacote que indica uma velocidade de *download* de 2MB. Vinte e cinco dos seus clientes mediram as suas velocidades de *download* num período que pode ser considerado de menor tráfego, registando-se uma velocidade média igual a 1.5Mb e um desvio padrão igual a 1 Mb. Haverá evidência, ao nível de 1%, para podermos concluir que a velocidade de *download* é inferior à indicada por esse fornecedor? Enuncie os pressupostos que entenda necessários.

(2Ch 2006-07-04)

3. Um determinado Município pretende efectuar uma sondagem junto da população que vive num bairro mais afastado a fim de determinar a proporção de pessoas que diariamente utilizam os transportes públicos.
- Determine o número de munícipes a inquirir de modo a obter um erro de estimativa máximo igual a 2% para um nível de confiança de 95%.
  - Os meios financeiros disponíveis apenas permitiam inquirir 1000 pessoas, das quais 150 afirmaram utilizar os transportes públicos regularmente. Obtenha um intervalo de confiança a 95% para a proporção de munícipes que utilizam regularmente os transportes públicos. Diga qual o erro de estimativa associado e comente o resultado.

3. Um construtor de pneus de automóveis está a testar a duração de dois tipos de pneus. Para isso, um pneu de cada tipo foi colocado em cada uma das rodas traseiras de oito automóveis de teste. Registou-se a duração (em quilómetros) de cada um dos pneus até que se desgastasse completamente:

Carro de teste		1	2	3	4	5	6	7	8
Tipo	Pneu1	36925	45300	36240	32100	37210	48360	38200	33500
pneu	Pneu2	34318	42280	35500	31950	38015	47800	37810	33215

Construa um intervalo de confiança de 99% para a diferença entre a duração média dos dois tipos de pneus. O que conclui?

3. Um engenheiro em informática está a investigar a influência de duas linguagens de programação na melhoria das tarefas de programação. A doze programadores experientes, familiarizados com ambas as linguagens, é pedido que codifiquem um dado algoritmo em ambas as linguagens, e o tempo em minutos é registado. A seguinte tabela resume os resultados obtidos:

Programador	Tempo (em minutos)	
	Linguagem 1	Linguagem 2
1	17	18
2	16	14
3	21	19
4	14	11
5	18	23
6	24	21
7	16	10
8	14	13
9	21	19
10	23	24
11	13	15
12	18	20

Encontre um intervalo de confiança de 95% para a diferença dos tempos médios de codificação nas duas linguagens. Há alguma evidência que uma das linguagens seja preferível?

3. Duas formulações de combustível estão a ser testadas. As variâncias do número de octanas de cada uma das formulações é  $\sigma_1^2 = 1.5$  e  $\sigma_2^2 = 1.2$ , respectivamente. Duas amostras aleatórias de tamanhos  $n_1 = 15$  e  $n_2 = 20$  são testadas, tendo se observado as seguintes médias:  $\bar{x}_1 = 88.85$  e  $\bar{x}_2 = 92.54$ . (Assuma a normalidade).
- Construa um intervalo de confiança de 95% para a diferença das médias do número de octanas das duas formulações de combustível.
  - Qual o tamanho das amostras requerido para que o erro da estimativa fosse inferior a 1, com 90% de confiança? (considere  $n_1 = n_2$ ).



3. Um construtor de pneus de automóveis está a testar a duração de dois tipos de pneus. Para isso, um pneu de cada tipo foi colocado em cada uma das rodas traseiras de oito automóveis de teste. Registou-se a duração (em quilómetros) de cada um dos pneus até que se desgastasse completamente:

Carro de teste		1	2	3	4	5	6	7	8
Tipo pneu	Pneu1	36925	45300	36240	32100	37210	48360	38200	33500
	Pneu2	34318	42280	35500	31950	38015	47800	37810	33215

Construa um intervalo de confiança de 99% para a diferença entre a duração média dos dois tipos de pneus. O que conclui?

3. Para decidir a compra de um novo *scanner* manual para registar os artigos de uma loja comercial, são feitos testes ao *scanner* actualmente em uso e ao novo modelo a adquirir. Para isso, mediu-se o número de códigos de barras que cada um dos *scanners* consegue ler por segundo, obtendo-se os seguintes resultados:

Novo <i>scanner</i>	<i>Scanner</i> actual
$n_1 = 61$	$n_2 = 61$
$\bar{x}_1 = 40$	$\bar{x}_2 = 29$
$s_1^2 = 24.9$	$s_2^2 = 22.7$

Será o desempenho médio do novo *scanner*, em termos de número de códigos barras lidos por segundo, superior ao actualmente utilizado? ( $\alpha = 0.05$ )

3. Um supervisor da produção suspeita que existe uma diferença entre as proporções de itens defeituosos produzidos por duas máquinas,  $p_1$  e  $p_2$ , respectivamente. De uma amostra de 100 itens produzidos por cada uma das máquinas, verificou-se que a máquina 1 produziu 4 itens defeituosos enquanto que a máquina 2 produziu apenas 2 itens defeituosos. Verifique se a máquina 1 produz mais defeitos do que a máquina 2. (Use  $\alpha = 0.05$ ).

3. Um estudo pretende determinar a quantidade de dinheiro gasto pelas famílias, no aquecimento, no último ano; do passado sabe-se que o desvio padrão dos gastos em aquecimento é de 130 unidades monetárias. Neste estudo é requerido que a estimativa do gasto médio em aquecimento, esteja correcta a menos de 10 unidades monetárias com 90% de confiança. Quantas famílias devem ser incluídas na amostra?

3. Comente, justificando, se cada uma das seguintes afirmações é correcta ou não:
- A probabilidade de que a hipótese nula seja verdadeira é igual a  $\alpha$ .
  - Se a hipótese nula é rejeitada, então o teste prova que a hipótese alternativa é verdadeira.
  - Em todos os testes estatísticos de hipóteses,  $\alpha + \beta = 1$ .

3. Considere a seguinte função de densidade de probabilidade,

$$f(x) = \lambda^2 x e^{-\lambda x} \quad \lambda > 0 \quad -\infty < x < \infty$$

Encontre o estimador de máxima verosimilhança para o parâmetro  $\lambda$ .

3. O diâmetro de um determinado cabo tem um desvio padrão de  $\sigma = 0.01$  cm. Uma amostra aleatória de dimensão 10 produziu um diâmetro médio de 1.5045 cm. Use  $\alpha = 0.01$ .
- Teste a hipótese nula de que o diâmetro é igual a 1.50cm contra a hipótese alternativa que o diâmetro é diferente de 1.50 cm.
  - Qual a probabilidade de observar uma valor médio igual ao da amostra?
  - Qual o tamanho da amostra necessário para detectar um verdadeiro diâmetro médio de 1.505 cm com uma probabilidade de pelo menos 0.9, ou seja, com um valor de  $\beta = 0.1$ ?
  - Qual o erro de tipo II se o verdadeiro diâmetro é de 1.505?
  - Encontre o intervalo de confiança de 99%. Parecem os resultados correctos à luz das respostas anteriores?

3. Duas máquinas são usadas para encher garrafas com 16 cl. O volume de enchimento pode ser assumido como normal, com um desvio padrão  $\sigma_1 = 0.041$  e  $\sigma_2 = 0.046$ . O controlador de qualidade suspeita que ambas as máquinas produzem o mesmo volume de enchimento, muito embora possa ser diferente de 16 cl. Uma amostra aleatória de 10 garrafas produziu os seguintes resultados:

$$\bar{x}_1 = 16.015$$

$$\bar{x}_2 = 16.005$$

- Pensa que o controlador de qualidade tem razão? Use  $\alpha = 0.05$ .
- Qual a probabilidade de observar uma diferença igual ou superior à observada?
- Qual a potência do teste se a verdadeira diferença das médias for igual a 0.06?
- Construa um intervalo de confiança de 95% e interprete o resultado.
- Assumindo amostras de igual dimensão, qual a dimensão da amostra necessária para assegurar que  $\beta = 0.05$  se a verdadeira diferença é 0.06, com  $\alpha = 0.05$ ?

4. O número de dias, numa semana, em que ocorreram acidentes de trabalho numa grande indústria, foi registado para uma amostra de 200 semanas. Verifique se os dados, apresentados se ajustam ao modelo Binomial com parâmetros  $n = 5$  e  $p = 0,2$  (use um nível de significância de ~~10%~~).

5%

Nº Acidentes	0	1	2	3	4	5
Frequência	64	56	40	24	8	8

(1Ch 2006-06-19)

4. Na tabela seguinte encontram-se listados os resultados de um inquérito realizado a 1000 famílias que procurou avaliar a existência associação entre a graduação dos pais e dos filhos em informática.

	Pais graduados em Informática	
	Sim	Não
Filhos graduados em Informática		
Sim	360	240
Não	160	240

Verifique se existe essa associação utilizando um nível de significância de 5%.

(2Ch 2006-07-04)

4. Um modelo de automóvel é vendido em quatro versões, SX, LX, GLX, GTX. Foi feita uma campanha publicitária para melhorar as vendas das versões GLX e GTX. Posteriormente, foi verificada a escolha das versões em 500 vendas escolhidas ao acaso. Os resultados foram: SX (205), LX (125), GLX (110), GTX (60). De acordo com o fabricante, a participação de cada versão nas vendas deste modelo até a realização da campanha era de 40% de SX, 30% de LX e 20% de GLX. Faça o teste estatístico adequado a 5% e verifique se há evidências de para afirmar que houve mudanças na participação das vendas de cada versão.

4. Um administrador hospitalar selecciona uma amostra aleatória de 1000 doentes para os quais o problema de saúde que os levou ao Serviços de Urgência é classificado em três estados de severidade, considerando também a idade dos doentes classificada em três grupos etários (ver tabela seguinte):

	Severidade do Problema		
Idade	Ligeira	Moderada	Grave
< 5 anos	220	60	40
5 – 14 anos	100	180	60
> 14 anos	120	100	120

Verifique se existe relação entre a severidade do problema e o escalão etário considerado (use  $\alpha = 0.05$ ).

4. Foi realizado um inquérito a 700 automobilistas acerca da concordância ou não com os limites de tolerância de álcool permitidos aos automobilistas. Com base nos resultados do inquérito, constantes na tabela, pretende-se verificar se a concordância acerca dos limites de álcool está relacionado com o sexo.

	Concorda	Não concorda
Mulheres	290	35
Homens	230	145

4. Os resultados abaixo provêm de um teste sorológico aplicado a indivíduos pertencentes a 3 amostras (A (70), B (40) e C (40)) compostas por indivíduos de provenientes de diferentes idades (crianças, adolescentes e adultos). Pretende-se verificar se a proporção de indivíduos com uma determinada reacção ao teste sorológico não difere significativamente nas três amostras.

AMOSTRA	Reacção +	Reacção -
A	25	45
B	15	25
C	10	30

4. Um administrador hospitalar selecciona uma amostra aleatória de 1000 doentes para os quais o problema de saúde que os levou ao Serviços de Urgência é classificado em três estados de severidade, considerando também a idade dos doentes classificada em três grupos etários (ver tabela seguinte):

Idade	Severidade do Problema		
	Ligeira	Moderada	Grave
< 5 anos	220	60	40
5 – 14 anos	100	180	60
> 14 anos	120	100	120

Verifique se existe relação entre a severidade do problema e o escalão etário considerado (use  $\alpha = 0.05$ ).

4. Num estudo sobre hábitos alimentares foram questionados 206 jovens que frequentam os 5º e 6º anos, de uma escola do distrito de Braga. O nutricionista responsável pelo estudo avaliou qual o peso dos jovens e questionou-os quanto ao hábito de adicionar molhos nas refeições do tipo maionese, *ketchup* e mostarda. Os resultados obtidos encontram-se na tabela seguinte:

Peso	molhos			Total
	sempre	às vezes	nunca	
<38	5	40	10	55
[38, 46[	5	59	15	79
[46, 54[	4	27	11	42
$\geq 54$	4	16	10	30
Total	18	142	46	206

Verifique se existe relação entre o peso e hábito de adicionar molhos nas refeições. Utilize  $\alpha = 0.05$ .

4. Num estudo sobre hábitos alimentares foram questionados 102 jovens do sexo masculino e 104 jovens do sexo feminino que frequentam os 5º e 6º anos, de uma escola do distrito de Braga. O nutricionista responsável pelo estudo procurou avaliar se a preferência pelo tipo de cozinhado (cozidos, grelhados ou fritos) era a mesma para os dois sexos. Os resultados obtidos encontram-se na tabela seguinte:

sexo	cozinhados			Total
	cozidos	grelhados	fritos	
masculino	32	26	44	102
feminino	40	28	36	104
Total	72	54	80	206

Que conclusão pode retirar do estudo? Utilize  $\alpha = 0.01$ .

4. Num estudo sobre hábitos alimentares foram questionados 206 jovens que frequentam os 5º e 6º anos, de uma escola do distrito de Braga. O nutricionista responsável pelo estudo avaliou qual o peso dos jovens e questionou-os quanto ao tipo de alimentos que comem frequentemente. Os resultados obtidos encontram-se na tabela seguinte:

Peso	o que come			Total
	sopa, peixe ou carne, fruta e legumes	peixe ou carne e farináceos	pizza ou sandes, bolos e chocolates	
<38	22	20	13	55
[38, 46[	35	27	17	79
[46, 54[	17	12	13	42
>= 54	9	11	10	30
<b>Total</b>	<b>83</b>	<b>70</b>	<b>53</b>	<b>206</b>

Que conclusão pode retirar do estudo? Utilize  $\alpha = 0.05$ .

4. Um estudo foi conduzido para determinar a incidência de tuberculose em relação aos grupos sanguíneos numa amostra de esquimós.

	Tipo de Sangue			
Severidade	O	A	AB	B
Moderada	7	5	3	13
Mínima	27	32	8	18
Ausente	55	50	7	24

O que pode concluir? (Utilize  $\alpha=0,05$ ).

4. Um estudo de mercado foi conduzido para estudar a relação da personalidade com a atitude relativa a carros pequenos. Uma amostra de 299 adultos numa área metropolitana foi questionada, sendo classificada em 3 grupos: cuidadosa, duvidosa e confiante. Posteriormente, os inquiridos foram questionados sobre a sua opinião sobre carros pequenos: favorável, neutra e desfavorável. Há alguma relação entre personalidade e atitude relativa a carros pequenos?

	Personalidade		
Atitude	Cuidadosa	Duvidosa	Confiante
Favorável	79	58	49
Neutra	10	8	9
Desfavorável	10	34	42

4. Uma grande firma electrónica emprega muitos trabalhadores com deficiência, e pretende determinar de que modo é que o tipo de deficiência poderá afectar o desempenho do trabalhador.

		Desempenho		
		Abaixo da média	Média	Acima da Média
Deficiência	Cegos	21	64	17
	Surdos	16	49	14
	Não Deficientes	29	93	28

Utilize um nível de significância  $\alpha = 0.05$  para decidir com base nos dados amostrais da tabela, se é razoável afirmar que a deficiência não afecta o desempenho do trabalhador.

4. Uma determinada marca de automóveis pretende avaliar se existem diferenças em 3 marcas de pneus quanto à durabilidade. Para tal foi conduzida uma experiência e registados os seguintes resultados:

		Pneu		
		Marca A	Marca B	Marca C
Durabilidade	< 32000 km	26	23	15
	32000 - 46000	118	93	116
	+ 46000 km	56	84	69

Utilize  $\alpha = 0.01$  para testar a hipótese de que a durabilidade do pneu não depende da marca do mesmo.

(1Ch 2007-06-19)

5. O director de uma escola decidiu investigar se as classificações obtidas pelos alunos de um determinado ano variavam significativamente de turma para turma e de disciplina para disciplina. Para o efeito, escolheu ao acaso 2 turmas e três disciplinas e recolheu, para as diferentes combinações turma-disciplina, amostras independentes de 4 classificações. Os resultados obtidos apresentam-se na tabela seguinte:

	Disciplina X	Disciplina Y	Disciplina Z
Turma A	14, 11, 12, 14	16, 15, 13, 13	10, 12, 14, 10
Turma B	10, 13, 12, 14	14, 16, 17, 12	11, 12, 14, 10

- a) Apresente a tabela ANOVA para estes dados, indicando as condições de aplicabilidade (Cálculos auxiliares  $\sum y_{ijk}^2 = 4071$   $STQ = 92,625$ )

- b) O que pode concluir acerca da experiência?

(1Ch 2006-06-19)

5. Um perfume pode ser produzido usando um de três compostos A e um de dois ingredientes B. A resposta é medida pela intensidade depois da aplicação do perfume. Os dados obtidos foram os seguintes:

	A1	A2	A3
B1	81; 70	103; 110	118; 138
B2	123; 131	142; 143	180; 189

- a) Apresente a tabela ANOVA para estes dados, indicando as condições de aplicabilidade. (Cálculos auxiliares  $\sum y_{ijk}^2 = 208162$   $STQ = 13596,67$ )

- b) O que pode concluir acerca da experiência?

(2Ch 2006-07-04)

5. A tensão à rotura (psi – pouns per square inch), de determinados componentes poliméricos, depende do tempo de polimerização e da temperatura do banho. Para investigar a relação entre estas duas variáveis foram realizadas experiências que produziram os seguintes dados:

Temperatura do Banho	Tempo de polimerização	
	20 min	60 min
100 °C	11.80; 11.75; 11.90	11.90; 12.00; 12.10
120 °C	10.55; 11.10; 11.20	9.90; 10.15; 9.40

- a) Apresente a tabela ANOVA para estes dados, indicando as condições de aplicabilidade. (Cálculos auxiliares:  $\sum y_{ijk}^2 = 1500,2775$   $STQ = 9,5223$ ).

- b) O que pode concluir acerca da experiência?



(RCh 2006-07-24)

5. O departamento de qualidade de uma empresa de refrigerantes inspeciona as grades de 24 garrafas para determinar se existem garrafas com um volume inferior ao especificado. A inspeção de 75 grades produziu os seguintes resultados, que representam o número de grades com garrafas com volume inferior ao especificado:

<b>Enchimento</b>	0	1	2	3
<b>Frequência</b>	39	23	12	1

Com base nestas 75 observações determine se a distribuição binomial se ajusta aos valores observados (use  $\alpha = 0.05$ ).

(1Ch 2005-06-23)

5. O departamento de qualidade de uma empresa de molas registou o número de defeitos em molas de aço galvanizado. Foram inspecionadas 75 molas com os seguintes resultados:

<b>Nº defeitos</b>	1	2	3	4	5	6	7	8
<b>Frequência</b>	1	11	8	13	11	12	10	9

Verifique se os dados seguem uma distribuição de Poisson. Use  $\alpha = 0.01$ .

(2Ch 2005-07-06)

5. Uma experiência foi realizada, no âmbito do controlo de qualidade, para investigar o efeito do tipo de vidro e do tipo de fósforo no brilho de ecrãs de televisão. A resposta é a corrente (em microamperes) necessários para obter um nível especificado de brilho. Os resultados obtidos foram os seguintes:

	<b>Fósforo 1</b>	<b>Fósforo 2</b>	<b>Fósforo 3</b>
<b>Vidro 1</b>	280;290;285	300;310;295	290;285;290
<b>Vidro 2</b>	230;235;240	260;240;235	220;225;230

Faça a análise apropriada explicitando todas as considerações subjacentes ao modelo. Use  $\alpha = 0.01$ . (Cálculos auxiliares:  $\sum_{i,j,k} y_{ijk}^2 = 1264350$ ; STQ = 16150).

(RCh 2005-07-26)

5. O departamento de qualidade de uma empresa de refrigerantes inspeciona as grades de 24 garrafas para determinar se existem garrafas com um volume inferior ao especificado. A inspeção de 75 grades produziu os seguintes resultados, que representam o número de grades com garrafas com volume inferior ao especificado:

<b>Enchimento</b>	0	1	2	3
<b>Frequência</b>	39	23	12	1

Com base nestas 75 observações determine se a distribuição binomial se ajusta aos valores observados (use  $\alpha = 0.05$ ).

5. Uma experiência foi realizada para investigar a deformação de uma placa de cobre. Os dois factores estudados foram a temperatura e quantidade de cobre das placas.

Na seguinte tabela estão as deformações da placa medidas na experiência realizada.

Temperatura (°C)	Quantidade de Cobre (%)			
	40	60	80	100
50	17, 20	16, 21	24, 22	28, 27
75	12, 9	18, 13	17, 12	27, 31
100	16, 12	18, 21	25, 23	30, 23
135	21, 17	23, 21	23, 22	29, 31

Existe alguma evidência de que a temperatura e a quantidade de cobre afectem a deformação da placa?

Haverá interacção entre os factores? ( $\alpha = 0.01$ ) (Dados auxiliares:  $\sum_{i,j,k} y_{ijk}^2 = 14703$  e  $SQR = 148.5$ ).

5. Para comparar as necessidades energéticas de três actividades físicas (correr, caminhar, andar de bicicleta), mediu-se o número de quilocalorias despendidas por quilómetro. Oito indivíduos foram escolhidos para realizar as actividades. A cada um deles foi atribuída aleatoriamente cada uma das actividades físicas (em termos de ordem). A todos os indivíduos, entre as actividades, foi concedido um período de descanso suficiente para evitar a fadiga. Cada actividade foi monitorizada exactamente uma vez para cada indivíduo, obtendo-se os seguintes resultados:

Indivíduo	Actividade Física		
	Correr	Caminhar	Andar de Bicicleta
1	1.4	1.1	0.7
2	1.5	1.2	0.8
3	1.8	1.3	0.7
4	1.7	1.3	0.8
5	1.6	0.7	0.1
6	1.5	1.2	0.7
7	1.7	1.1	0.4
8	2.0	1.3	0.6

Supondo que o número de quilocalorias despendidas por quilómetro é uma variável normalmente distribuída, teste a possibilidade de haver diferenças no número de quilocalorias despendidas por quilómetro entre as actividades físicas. Haverá diferenças metabólicas entre os indivíduos? Utilize  $\alpha = 0.05$ . (Dados auxiliares:

$$\sum_{i,j} y_{ij}^2 = 36.18 \text{ e } STQ = 5.353).$$

5. A decomposição das folhas de uma determinada árvore em diferentes condições ambientais foi estudada para três períodos de tempo de exposição distintos. Pacotes idênticos de folhas foram expostos a quatro ambientes distintos durante os períodos de tempo considerados. No final, mediu-se a diminuição do peso de cada um dos pacotes (em gramas), obtendo-se os seguintes resultados:

Ambiente	Tempo de exposição		
	1 mês	2 meses	3 meses
1	1.09, 1.06	1.35, 1.53	1.60, 1.40
2	1.16, 1.03	1.38, 1.35	2.18, 1.77
3	1.01, 1.04	1.63, 1.51	1.66, 1.98
4	0.90, 1.03	1.60, 1.72	1.73, 1.76

Existe alguma evidência de que as diferentes condições ambientais e os diferentes períodos de tempo de exposição afectem a decomposição das folhas? Haverá interacção entre os factores? ( $\alpha = 0.05$ ) (Dados

auxiliares:  $\sum_{ijk} y_{ijk}^2 = 52,2083$  e  $SQR = 0,20455$ ).

(1Ch 2003-06-12)

5. Os dados representam os pesos de alimentos consumidos (em Kg) consumidos por veados adultos em diferentes meses do ano.

	Fevereiro	Maior	Agosto	Novembro
	4.7	4.6	4.8	4.9
	4.9	4.4	4.7	5.2
	5.0	4.3	4.6	5.4
	4.8	4.4	4.4	5.1
	4.7	4.1	4.7	5.6
		4.2	4.8	

- a) Teste a hipótese que o consumo de alimentos é o mesmo para todos os meses testados. Utilize a ANOVA, considerando  $\alpha=0,05$ .
- b) Como faria a análise dos resíduos?

(2Ch 2003-06-30)

5. Um laboratório possui quatro técnicos para determinar o teor de fósforo em cereais, que se supõe ser uma variável normalmente distribuída. Aos diversos técnicos foram aleatoriamente distribuídas amostras de um mesmo cereal. Os resultados obtidos foram os seguintes:

T1	T2	T3	T4
34	37	34	36
36	36	37	34
34	35	35	37
35	37	37	34
34	37	36	35

- a) Verifique se existem diferenças entre os técnicos que executam a análise.
- b) Como faria a análise dos resíduos?

5. Uma empresa de recrutamento desenvolveu um teste para determinar a capacidade de aprendizagem dos candidatos finais a uma vaga de Director de Investigação e Desenvolvimento (I&D) numa multinacional. O teste consistia em três listas com 20 palavras cada, dispondo cada candidato de 5 minutos para estudar cada uma das listas e memorizar o maior número possível de palavras. Os resultados obtidos pelos candidatos foram os seguintes:

	Lista 1	Lista 2	Lista 3
Ana	18	14	16
António	15	14	13
Carlos	12	11	11
Marisa	10	11	12

Verifique se existem diferenças significativas no grau de dificuldade médio entre as 3 listas de palavras (use  $\alpha=0.1$ ).

5. Os dados representam o tempo de atendimento (min), numa estação de correios. Verifique se os dados se ajustam a uma distribuição exponencial.

Intervalo	[0-4[	[4-8[	[8-12[	[12-16[	[16-20[	[20-24[	[24-28[	[28-32[	$\geq 32$
Observações	125	98	43	21	12	8	2	1	0

(1Ch 2007-06-19)

6. Uma amostra aleatória refere-se a uma certa pontuação de um teste psicológico para verificar o nível de destreza entre dois grupos formados por 10 meninas e 10 meninos de 5 anos de idade.

Meninas	15	16	10	14	18	16	15	25	20	21	$\bar{x}_1 = 17$	$s_1 = 3,97$
Meninos	11	14	16	16	18	19	22	27	25	27	$\bar{x}_2 = 19,5$	$s_2 = 5,28$

Verifique se a distribuição de pontuações no teste é a mesma para as meninas e meninos. Considere  $\alpha = 5\%$ .

(1Ch 2006-06-19)

6. Uma determinada empresa com representações no Porto e em Lisboa decide promover um funcionário para um cargo de chefia. No processo de selecção do funcionário a empresa pretende, antes de mais, averiguar se existem diferenças entre a competência dos funcionários do Porto e a competência dos funcionários de Lisboa. Para tal foram seleccionados ao acaso 10 funcionários de Lisboa e 9 do Porto, tendo estes sido sujeitos a uma prova cotada de 0 a 20. Os resultados obtidos nas duas cidades foram os seguintes:

<b>Lisboa</b>	13,6	16,8	3,0	15,5	12,5	12,2	18,5	12,8	18	16,7
<b>Porto</b>	13,4	19,0	15,7	16,0	18,2	12,3	2,0	14,4	14,6	

Efectuando o teste mais apropriado, qual a decisão que este lhe permite chegar utilizando um nível de significância de 5%.

(2Ch 2006-07-04)

6. Os tempos entre as chegadas de nove fregueses consecutivos ao balcão de informações de uma loja foram, em minutos,

1,0    0,6    1,2    0,6    1,8    0,3    0,5    2,0

Teste a hipótese de que os tempos entre as chegadas obedecem a uma distribuição exponencial. (Use  $\alpha = 1\%$ )

(RCh 2006-07-24)

6. Uma experiência foi realizada para determinar se a temperatura de ignição ou a posição da fornalha afectam a densidade de um ânodo de carbono. Os resultados obtidos foram os seguintes:

	Temp 800°C	Temp 825°C	Temp 850°C
<b>Posição 1</b>	570;565;583	1063;1080;1043	565;510;590
<b>Posição 2</b>	528;547;521	988;1026;1004	526;538;532

Faça a análise apropriada explicitando todas as considerações subjacentes ao modelo. Use  $\alpha = 0.01$ . (Cálculos auxiliares:  $\sum_{i,j,k} y_{ijk}^2 = 10031071$ ; STQ = 958690,944).

(1Ch 2005-06-23)

6. Um engenheiro quer verificar se o acabamento de superfícies metálicas é influenciado pelo tipo de tinta usada e o tempo de secagem. Para o efeito, seleccionou 3 tempos de secagem, 20, 25 e 30 minutos e usou 2 tipos de tinta. Três peças foram testadas em cada combinação de tipo de tinta e tempo de secagem. Os resultados obtidos foram os seguintes:

	Tempo 20min	Tempo 25min	Tempo 30min
Tinta 1	74;64;50	73;61;44	78;85;92
Tinta 2	92;86;68	98;73;88	66;45;85

Faça a análise apropriada explicitando todas as considerações subjacentes ao modelo. Use  $\alpha = 0.05$ .

(Cálculos auxiliares:  $\sum_{i,j,k} y_{ijk}^2 = 101598$ ; STQ = 4504.44)

(2Ch 2005-07-06)

6. Os dados dizem respeito a uma amostra aleatória de embalagens de 1000gr retiradas de uma máquina automática de embalar.

983 992 1022 976 997 1000 1004 983 998

Verifique se os dados seguem uma distribuição Normal (use  $\alpha = 0.01$ ).

(RCh 2005-07-26)

6. Uma experiência foi realizada para determinar se a temperatura de ignição ou a posição da fornalha afectam a densidade de um ânodo de carbono. Os resultados obtidos foram os seguintes:

	Temp 800°C	Temp 825°C	Temp 850°C
Posição 1	570;565;583	1063;1080;1043	565;510;590
Posição 2	528;547;521	988;1026;1004	526;538;532

Faça a análise apropriada explicitando todas as considerações subjacentes ao modelo. Use  $\alpha = 0.01$ . (Cálculos auxiliares:  $\sum_{i,j,k} y_{ijk}^2 = 10031071$ ; STQ = 958690,944).

(1Ch 2004-06-26)

6. Pretende-se estudar o número de defeitos ocorridos no fabrico de placas de circuito impressas. Uma amostra aleatória de 60 placas de circuito impressas foi considerada e o número de defeitos registado, obtendo-se os seguintes dados:

Número de defeitos	Número de placas
0	32
1	15
2	9
3	4

Será que se pode afirmar que o número de defeitos ocorridos no fabrico de placas de circuito impressas segue uma distribuição de Poisson? ( $\alpha = 0.05$ )

(2Ch 2004-07-09)

6. Para avaliar o desempenho de um programa de treino num dado sistema organizacional seleccionaram-se aleatoriamente 18 gestores de quadros médios, os quais foram distribuídos por 3 grupos. Ao grupo A não foi ministrado nenhum programa de treino, ao grupo B foi ministrado um programa com assistência por computador e ao grupo C foi ministrado um programa com assistência por computador conjuntamente com umas técnicas comportamentais. Após o período de treino os gestores foram sujeitos a um exame com 25 questões de escolha múltipla, registando-se o número de respostas correctas para cada um. Os resultados obtidos foram:

Sem treino	Treino assistido por computador	Treino assistido por computador + técnicas
16	19	12
18	22	19
11	13	18
14	15	22
23	20	16
	18	25
	21	

Existe evidência suficiente para indicar que a distribuição de classificações difere em média para os três programas considerados? Utilize  $\alpha = 0.01$ .

(RCh 2004-07-24)

6. No processo de fabrico de componentes electrónicos, estes são submetidos a um banho de galvanização. Realizou-se a medição do pH de 10 amostras retiradas do banho de galvanização, obtendo-se os seguintes valores:

7.91    7.85    6.82    8.01    7.46    6.95    7.05    7.35    7.25    7.42

Os engenheiros industriais afirmam que o pH do banho de galvanização segue uma distribuição Normal com média 7.4 e desvio padrão 0.4. Parecem os dados indicar que esta afirmação é correcta? ( $\alpha=0.05$ )

(1Ch 2003-06-12)

6. Os dados representam os pesos em Kg de duas amostras de estudantes de Braga e de Faro. Verifique se poderemos considerar as distribuições dos pesos idênticas. Considere  $\alpha=0,1$

87	75	77	84	55	67	84	93	70	82	76	52	74	66	71
94	79	80	46	46	84	89	72	55	79	67	66	70	58	48

(2Ch 2003-06-30)

6. Um estudo preliminar foi conduzido para determinar os níveis de base de uma substância tóxica PCB (policlorinato bifenílico) em amostras de solo num determinado país. A tabela apresenta os níveis de PCB (0.0001 g/kg de solo) em amostras de solos recolhidas em 10 localizações rurais e 10 urbanas.

Rural		Urbana	
8.1	9.8	24.0	11.0
18.0	15.0	16.0	22.0
19.0	12.0	21.0	13.0
12.0	11.0	94.0	12.0
16.0	23.0	11.0	18.0

Poder-se-á concluir que as distribuições dos níveis de PCB nas duas áreas são diferentes?

(1Ch 2002-06-03)

6. A empresa “BRANCURATOTAL” está interessada em avaliar o desempenho de três marcas diferentes de detergentes. Foi efectuado um teste de brancura em três marcas de máquinas diferentes e repetida a experiência 2 vezes para cada conjunto detergente/máquina de lavar. Os resultados obtidos foram os seguintes:

		Máquina					
		A		B		C	
Detergente	1	45	39	43	42	51	58
	2	47	44	46	46	52	48
	3	48	34	50	47	55	45

Que pode concluir sobre a interacção entre o tipo de detergente e a máquina de lavar? Utilize  $\alpha = 0.05$ . [Dados Auxiliares:  $\sum y_{ijk}^2 = 39728$  ;  $SQR = 208$  ]

(2Ch 2002-06-20)

6. Para testar o desempenho de uma raquete de ténis, pediu-se que um jogador jogasse com uma raquete durante uma hora e a classificasse numa escala de 1 a 10. Foram avaliadas duas características da raquete,: a dimensão (padrão e grande) e a tensão dos fios (baixa, média e elevada). As raquetes foram distribuídas aleatoriamente por oito jogadores e registaram-se os seguintes resultados:

		Tensão do fio								
		baixa			média			elevada		
Dimensão	Padrão	5	6	5	6	6	7	5	4	6
	Grande	6	7	5	8	9	7	3	5	2

Efectue a análise da variância apropriada e indique em que condições o modelo é válido. Utilize  $\alpha = 0.01$ . [Dados Auxiliares:  $\sum y_{ijk}^2 = 626$  ;  $SQR = 12$  ].



(1Ch 2007-06-19)

7. Suponha que se pretende determinar uma relação entre a percentagem de partos assistidos por pessoal de saúde especializado e a taxa de mortalidade por 100 000 habitantes. Os dados para uma amostra de cinco países são apresentados na tabela:

País	% assistida (X)	tx. Mortalidade (Y)
Hungria	95	26
Irão	82	120
Quênia	28	250
Paquistão	24	400
Portugal	87	12

- a) Determine a recta de mínimos quadrados.  
 b) Verifique se existe relação linear estatisticamente significativa.  
 c) Qual a taxa de mortalidade por 100 000 habitantes, para um país cujo registo de percentagem de partos assistidos por pessoal de saúde especializado é de 90%. E para 20%?

(1Ch 2006-06-19)

7. Um criador de gado ovino afirma que o peso das ovelhas existentes nos rebanhos de uma dada zona de pastorícia segue uma distribuição Normal. Recolheu-se uma amostra de 10 ovelhas, cujos pesos (em kg) foram

65 75 75 70 75 70 80 70 75 80

Com base nestes dados o que poderá dizer quanto à afirmação do criador? (Utilize  $\alpha = 5\%$ )

(2Ch 2006-07-04)

7. Uma companhia de seguros de automóveis classifica seus segurados em fumadores e não-fumadores. Para oito grupos, de 100 fumadores cada um, a frequência de acidentes foi a seguinte:

7 9 5 13 8 11 8 6

As frequências correspondentes para doze grupos de não-fumadores foram:

4 2 4 3 3 6 3 4 2 4 3 2

Verifique se esses dados confirmam a hipótese de que a distribuição de acidentes dos não fumadores é significativamente menor do que o dos fumadores.

(RCh 2006-07-24)

7. Os dados dizem respeito a uma amostra aleatória de embalagens de bolos produzidos por uma dada padaria.

983 995 1022 976 995 1000 1004 983 998

Verifique se os dados seguem uma distribuição Normal com média 1000 gr e desvio padrão 15 gr. (Use  $\alpha = 0.1$ ).

(1Ch 2005-06-23)

7. Um engenheiro electrotécnico pretende estudar distribuição da corrente fornecida por dois tubos catódicos. Para o efeito, seleccionou amostras de cada tipo de tubo. Os dados obtidos foram os seguintes (micro amperes):

<b>Tubo A</b>	251	255	258	257	250	251	254	250	248
<b>Tubo B</b>	250	253	249	256	259	252	260	251	

Verifique se as distribuições de corrente para os dois tubos são diferentes.

(2Ch 2005-07-06)

7. Um engenheiro em informática está a investigar a distribuição dos tempos de codificação de um dado algoritmo utilizando duas linguagens de programação diferentes. A vinte programadores experientes, familiarizados com ambas as linguagens, é pedido que codifiquem um dado algoritmo numa das linguagens, e o tempo em minutos é registado. A seguinte tabela resume os resultados obtidos:

<b>Tempo (em minutos)</b>	
<b>Linguagem 1</b>	<b>Linguagem 2</b>
17	18
16	14
21	19
14	11
18	23
24	21
16	10
14	13
	19
	25

Verifique se as distribuições dos tempos de programação para as duas linguagens são idênticas. Use  $\alpha = 0.1$ .

(RCh 2005-07-26)

7. Os dados dizem respeito a uma amostra aleatória de embalagens de bolos produzidos por uma dada padaria.

983   995   1022   976   995   1000   1004   983   998

Verifique se os dados seguem uma distribuição Normal com média 1000 gr e desvio padrão 15 gr. (Use  $\alpha = 0.1$ ).

(1Ch 2004-06-26)

7. Smart modems foram construídos com microprocessadores com características avançadas que garantem o desempenho e a eficiência global de um modem. Três modems com microprocessadores próprios existem no mercado: *Bizcomp 1012*, *Hayes Smartmodem 1200* e *Vadic 3451 Auto-Dial*. Suponha que 5 utilizadores de cada tipo de modem foram aleatoriamente seleccionados e classificaram os modems quanto ao desempenho (numa escala de 1 a 100). Baseando-se nos dados que se encontram na tabela, que pode concluir quanto ao desempenho dos diferentes modems? Utilize  $\alpha = 0.05$ .

	<i>Bizcomp 1012</i>	<i>Hayes Smartmodem 1200</i>	<i>Vadic 3451 Auto-Dial</i>
	63	69	78
	72	70	86
	80	72	90
	87	80	94
	91	83	98
médias	78,60	74,80	89,20
desvios padrão	11,33	6,30	7,69

(2Ch 2004-07-09)

7. As *motherboards* de um certo fabricante possuem uma bateria para manter a configuração do computador. Estas baterias devem ser substituídas ocasionalmente. Em 10 *motherboards* deste fabricante, mediu-se o tempo (em anos), até que fosse necessária a substituição das respectivas baterias, obtendo-se os seguintes resultados:

1.7 4.0 1.9 2.0 1.7 2.1 2.7 4.2 1.8 2.2

Verifique se a duração das baterias das *motherboards* deste fabricante pode ser considerada exponencial ( $\alpha = 0.05$ ).

(RCh 2004-07-24)

7. A Comissão Federal de Comunicações (CFC) especifica que as emissões de radiação electromagnética de dispositivos digitais devem ser medidas num local teste em campo aberto. Para verificar a aceitabilidade desse local teste, a designada atenuação do local deverá ser avaliada. Um estudo conduzido num local teste nos EUA em 1985, conduziu aos seguintes resultados:

Frequência de transmissão, $X$ (MHz)	50	100	200	300	400	500	600	700	800	900	1000
Atenuação do local, $Y$ (db)	11,5	15,8	18,2	22,6	26,2	27,1	29,5	30,7	31,3	32,6	34,9

- a) Determine o coeficiente de correlação de Spearman entre a frequência de transmissão ( $X$ ), e a atenuação ( $Y$ ). Interprete o seu valor.
- b) Será que um maior valor de atenuação corresponderá a uma transmissão de frequência mais elevada? Efectue o teste estatístico adequado, utilizando  $\alpha = 0.10$ .

(1Ch 2003-06-12)

7. Os dados representam o peso, em gramas, de 12 caranguejos. Verifique se a distribuição dos pesos segue uma distribuição Normal com média 11 gramas e desvio padrão de 5 gramas. Utilize  $\alpha=0,01$

15.39      14.90      4.19      17.25      11.30      22.70  
2.50      15.20      1.41      15.81      9.52      14.40

(2Ch 2003-06-30)

7. Os dados representam o peso, em gramas, de 12 caranguejos. Verifique se a distribuição dos pesos segue uma distribuição normal.

15.39	14.90	4.19	17.25	11.30	22.70
2.50	15.20	1.41	15.81	9.52	14.40

(1Ch 2002-06-03)

7. Um estagiário da Universidade do Minho foi encarregue de analisar o processo de embalagem de uma máquina de embalagem de pacotes de 500 g de rebuçados. Para tal retirou uma amostra aleatória de 10 pacotes e obteve os seguintes resultados:

489.5   490.0   490.0   498.5   499.0   501.5   502.5   502.5   510.0   510.5

Suponha que na revisão periódica a máquina tinha sido calibrada para que o peso dos pacotes de rebuçados fosse normalmente distribuído, com média de 500 g e desvio padrão de 5.1 g. Perante a amostra recolhida, o estagiário pode afirmar que estas normas estão a ser respeitadas?

(2Ch 2002-06-20)

7. Um canal de televisão privado decidiu fazer um estudo que incluía a exibição de um filme dobrado em horário nobre e, uma sondagem que compreendia a questão: “*Concorda que os filmes estrangeiros exibidos na televisão devam ser dobrados em português?*”. As respostas obtidas uma semana antes e uma semana depois da exibição do filme, foram codificadas de 1 (discordo totalmente) a 10 (concordo totalmente):

semana -1	6	8	6	10	1	8
semana +1	10	5	7	4	2	10

Verifique se existem diferenças na distribuição das respostas antes e depois da exibição do filme dobrado (use  $\alpha=0.1$ ).

(1Ch 2007-06-19)

8. Numa Universidade foram feitas as médias de aprovação no primeiro semestre dos doze alunos classificados à entrada num curso de Medicina e cujos resultados foram:

Classificação à entrada	1°	2°	3°	4°	5°	6°	7°	8°	9°	10°	11°	12°
Média da aprovação na Universidade	19,0	17,4	18,0	17,2	18,6	18,8	16,8	17,6	17,0	14,0	14,6	14,2

Verifique se as médias de aprovação na Universidade estão associadas com a classificação à entrada.

(1Ch 2006-06-19)

8. A tabela apresenta a relação entre as temperaturas máximas do dia e a procura máxima de electricidade, respectivamente. Pretende-se investigar se há uma relação linear entre o pico de procura de electricidade e a temperatura máxima diária observada.

$X (^{\circ}\text{C})$	33	29	35	39	31	36
$Y (\text{MW})$	207	139	211	273	156	244

- a) Ajuste um modelo do tipo

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

- b) Teste a hipótese de que não existe relação entre o pico de procura de electricidade e a temperatura máxima diária observada. O que pode concluir?

(2Ch 2006-07-04)

8. Os dados representam o rendimento per capita (milhares de dólares) em função do número de habitantes de alguns países (em milhões).

Pop X	54	42	28	38	25	70	48	41	20	52	65
Rend Y	6	16	33	18	41	3	10	14	45	9	5

- a) Ajuste um modelo do tipo

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 \frac{1}{X_i}$$

- b) Teste a hipótese de que não existe relação entre o rendimento e a população. O que pode concluir? (Cálculos auxiliares:  $\hat{\sigma}^2 = 6,469$ ).

8. A tabela seguinte indica as idades e as pressões arteriais sistólicas (PAS) de 12 mulheres seleccionadas aleatoriamente de um conjunto de mulheres.

Idade , X	Pressão arterial, Y
56	147
42	125
72	160
36	118
63	149
47	128
55	150
49	145
38	115
42	140
68	152
60	155

Verifique se existe associação entre a idade e a PAS (use  $\alpha = 0.05$ ).

8. Suspeita-se que um bom aluno a Matemática produza bons resultados a Estatística. Foram seleccionado 10 alunos de um curso, registando-se as classificações obtidas nas duas disciplinas num teste final numa escala de 0 -100.

$X_i$ (Matemática)	45	65	72	71	31	50	65	40	80	61
$Y_i$ (Estatística.)	73	50	67	75	44	42	72	48	50	62

Verifique se existe associação entre os resultados obtidos nas duas disciplinas (use  $\alpha = 0.05$ ).

8. Pensa-se existir uma relação entre o número de octanas de uma gasolina e a potência debitada por certo tipo de motor a 5000 rotações por minuto. Um estudo efectuado conduziu aos seguintes resultados:

$X_i$	90	92	94	95	96	98	100
$Y_i$	83	83	87	92	93	94	94

Verifique se à medida que aumenta o número de octanas da gasolina a potência debitada pelo motor também aumenta (use  $\alpha = 0.05$ ).

8. A tabela seguinte indica as idades e as pressões arteriais sistólicas (PAS) de 12 mulheres seleccionadas aleatoriamente de um conjunto de mulheres.

Idade , X	Pressão arterial, Y
56	147
42	125
72	160
36	118
63	149
47	128
55	150
49	145
38	115
42	140
68	152
60	155

Verifique se existe associação entre a idade e a PAS (use  $\alpha = 0.05$ ).

8. Os dados representam o lucro  $Y$ , das vendas de uma companhia de construção, em nove projectos diferentes, e o número de anos de experiência  $X$ , do chefe da obra responsável pelo projecto.

$X$	2.5	3.5	4	5	6	6.5	7
$Y$	5	6	10	11	13	15	17

- Encontre a recta de mínimos quadrados que melhor se ajusta aos dados.
- Teste a hipótese que não há relação linear entre as duas variáveis, isto é, que o declive é nulo? O que conclui?
- Quais as condições necessárias para que as conclusões do modelo sejam válidas?
- Num qualquer modelo de regressão linear simples, qual o valor do declive que indica que  $X$  e  $Y$  não estão relacionados?

8. Os dados da tabela representam o crescimento dum organismo celular em circunstâncias controladas.

<b>Semanas</b>	1	3	4	5	6	8
<b>Altura</b>	0.5	2	8	12.5	18	31

- Encontre a recta de mínimos quadrados que melhor se ajusta aos dados.
- Teste a hipótese que não há relação linear entre as duas variáveis, isto é, que o declive é nulo? O que concluir?

(1Ch 2003-06-12)

8. A distância ( $y$ ) até à imobilização de um automóvel numa certa estrada foi estudada em função da velocidade.

Vel. (m/h)	Dist.(ft)
20.5	15.4
20.5	13.3
30.5	33.9
40.5	73.1
48.8	113.0
57.8	142.6

a) Ajuste um modelo linear de  $y$  em função da velocidade.

b) Ajuste um modelo linear  $\sqrt{y}$  em função da velocidade.

c) Qual o melhor modelo? Justifique. 
$$r^2 = \frac{\beta^2 \sum_i (X_i - \bar{X})^2}{\sum_i (Y_i - \bar{Y})^2}$$

(2Ch 2003-06-30)

8. A tabela apresenta os comprimentos de 13 andorinhas em função de várias idades.

Idade (dias), X	3	4	5	6	8	9	10	11	12	14	15	16	17
Comp. (cm), Y	1.4	1.5	2.2	2.4	3.1	3.2	3.2	3.9	4.1	4.7	4.5	5.2	5.0

a) Ajuste um modelo linear.

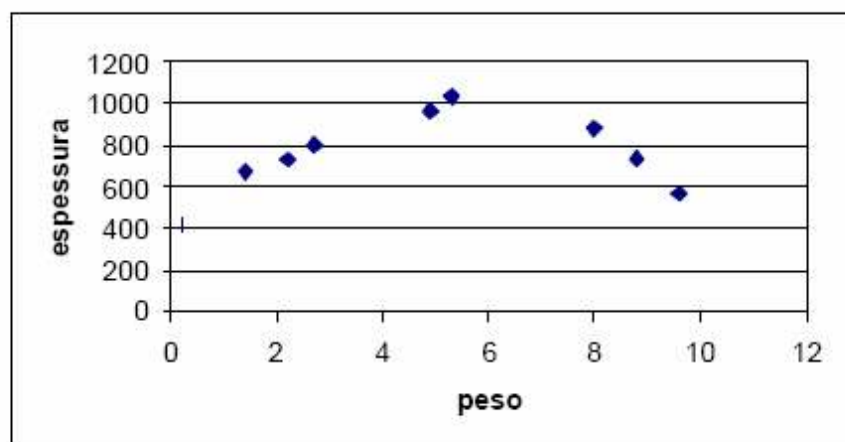
b) Calcule o intervalo de confiança para o declive.



(1Ch 2002-06-03)

8. A tabela apresenta o peso do olho em gramas e a espessura da córnea em micrómetros para nove bovinos. Verifique que o modelo de regressão de Y em função de X é quadrático contra a hipótese que é linear.

Peso (X)	0.2	1.4	2.2	2.7	4.9	5.3	8.0	8.8	9.6
Espessura (Y)	416	673	733	801	967	1036	883	736	567



Dados auxiliares:

$$\sum_i X_i = 43.1$$

$$\sum_i Y_i = 6812$$

$$\sum_i x_i^2 = 93.429$$

$$\sum_i z_i^2 = 10043.69$$

$$\sum_i Y_i X_i = 1391.889$$

$$\sum_i Y_i z_i = 1866.204$$

$$\sum_i x_i z_i = 941.964$$

Onde:

$$x_i = X_i - \bar{X}$$

$$Z_i = X_i^2$$

$$z_i = Z_i - \bar{Z}$$

(2Ch 2002-06-20)

8. O decréscimo da tensão superficial do cobre líquido (Y) parece ser função do logaritmo da percentagem de enxofre (X).

- Encontre a recta de regressão em função de X.
- Encontre a recta de regressão em função do logaritmo de X.
- Com base no coeficiente de determinação, qual o melhor modelo?

Enxofre %	Ln Enxofre %	Decréscimo (deg/cm)
0,034	-3,38	308
0,093	-2,38	426
0,301	-1,2	590
0,399	-0,92	624
0,613	-0,49	649
0,827	-0,19	727

Dados auxiliares:

$$\sum_i X_i = 2.267$$

$$\sum_i Z_i = -8.56$$

$$\sum_i Y_i = 3324$$

$$\sum_i Y_i X_i = 1475.722$$

$$\sum_i Y_i Z_i = -3793.14$$

$$\sum_i X_i^2 = 1.3193$$

$$\sum_i Z_i^2 = 19.6514$$

Onde:

$$Z_i = \ln X_i.$$

Exame	1	2	3	4	5	6	7	8
3-Jun-02								
20-Jun-02								
12-Jun-03								
30-Jun-03								
26-Jun-04								
09-Jul-04								
24-Jul-04								
23-Jun-05								
06-Jul-05								
26-Jun-05								
19-Jun-06								
04-Jul-06								
24-Jul-06								
19-Jun-07								