

1. Numa fábrica de automóveis existe uma secção destinada à produção de determinado tipo de peças, cujo comprimento deverá ser aproximadamente de 4.5 cm. A secção de controlo de qualidade da referida fábrica afirma que as peças apresentam comprimentos superiores aos exigidos. Com o objetivo de avaliar a veracidade da afirmação proferida pela secção de controlo de qualidade, selecionou-se ao acaso uma amostra de 20 peças na produção de um dia. A média da amostra foi 5.07 cm e o respetivo desvio padrão 0.776 cm. Admita que a população dos comprimentos é bem modelada por uma distribuição Normal.
 - 1.1 Determine estimativas intervalares a 90%, 95% e 99% de confiança para o comprimento médio das peças. Comente os resultados obtidos quanto à precisão e confiança.
 - 1.2 Indique uma estimativa pontual para o comprimento médio das peças.
 - 1.3 Se usar como estimativa do comprimento médio das peças o valor de 5.07 cm qual o erro máximo que está a cometer. Responda a esta questão com uma confiança de 95%.
 - 1.4 Se para a estimativa do comprimento médio das peças pretender um erro máximo de 0,2 cm com uma confiança de 95%, o que terá que fazer?
 - 1.5 Teste, ao nível de significância de 5%, se o comprimento médio das peças é superior a 4,5 cm. Confronte os resultados obtidos com os resultados obtidos na alínea 3.1.
 - 1.6 Teste, ao nível de significância 5%, se a variância do comprimento das peças é significativamente diferente de 1.
 - 1.7 Confirme o resultado anterior através de um intervalo de confiança. Use $\lambda=0.95$.
2. Estudos anteriores levam a supor que, quando uma criança de 2 meses começa a tomar exclusivamente leite do tipo A, o seu peso sofre um aumento que no primeiro mês é suposto seguir uma distribuição Normal de variância 9000 gramas². Foram escolhidas, ao acaso, 20 crianças de 2 meses a quem se deu o leite referido e o seu peso aumentou em média 475 gramas. Averigue, ao nível de significância 5% se o aumento de peso foi significativamente inferior a 500 gramas.
 - 2.1 Efetue o teste em causa com base no cálculo da região crítica.
 - 2.2 Confirme o resultado da alínea anterior através do cálculo de um intervalo de confiança a 95%.
3. O consumo mensal de calorias (kcal/g) duma certa espécie de esquilos é bem modelado por uma distribuição Normal. Recolheu-se uma amostra aleatória de dimensão 100 cuja média amostral foi de 2.3 e o desvio padrão amostral foi de 0.14.
 - 3.1 Teste, ao nível de significância 1%, se o consumo médio superior a 2.2.
 - 3.2 Confirme o resultado da alínea anterior através do cálculo de um intervalo de confiança para a média a 99%.
 - 3.3 Qual deve ser o tamanho da amostra a recolher para que, com 99% de confiança, o erro máximo cometido seja inferior a 0.05.

4. O ficheiro *ficha4.sav* contém uma amostra de valores correspondentes ao comprimento, em mm, de 60 parafusos produzidos pela secção A de uma fábrica e selecionados aleatoriamente. A média da amostra é 5.315 e o respetivo desvio padrão é 0.8293.

4.1 Averigüe se a média é diferente de 5.15, tendo por base o p-value do teste de hipóteses adequado e considerando um nível de significância de 10%

4.2 Efetue um teste para averiguar se a média é significativamente superior a 5.15 ($\alpha = 0.1$).

4.3 Obtiveram-se os seguintes resultados em SPSS:

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
comprimento de um parafuso	60	5,315	,8293	,1071

One-Sample Test

	Test Value = 5.15					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
comprimento de um parafuso	1,541	59	,129	,1650	-,049	,379

- Indique o significado de cada um dos campos das tabelas anteriores.
- Através da análise do p-value da tabela de output do teste t efetuado pelo SPSS confirme a resposta dada na alínea 6.1).
- Através da análise do p-value da tabela de output do teste t efetuado pelo SPSS confirme a resposta dada na alínea 6.2).

4.4 Confirme o resultado da alínea i) através do intervalo de confiança.

5. O ficheiro *ficha4.sav* contém medições do diâmetro, altura e volume de um conjunto de árvores. Considere as variáveis altura e volume. Utilize o SPSS para responder às seguintes questões:

5.1 Recorrendo ao SPSS efetue um teste para averiguar se a altura média da população de onde foram retirados os dados se pode considerar como sendo 70 ($\alpha = 0.01$).

- Indique as hipóteses em teste;
- Registe o valor do p-value e do valor observado da estatística de teste;
- Conclua.
- Confirme o resultado da alínea anterior através do cálculo de um intervalo de confiança para a média a 99%.

5.2 Teste, ao nível de significância 1%, se o desvio padrão da população de alturas é superior a 5.

6. "O estudo da Quercus, com a colaboração da Universidade da Madeira, revela que o pagamento pelos sacos de plástico aumenta 50% a sua taxa de reutilização e otimiza o seu uso em 20%. A Quercus defende uma taxa obrigatória sobre cada saco." (<http://naturlink.sapo.pt/article.aspx?menuid=20&cid=4946&bl=1>). A análise dos dados revela

que nos supermercados onde os sacos são oferecidos a utilização de sacos novos ocorre em 95% dos clientes (Intervalo de Confiança: 93%-97%), uma taxa muito maior do que aquela que ocorre nos supermercados onde os sacos são pagos (51%, Intervalo de Confiança: 47%-55%), revelando-se o pagamento dos sacos um forte contributo para a redução na produção de sacos de plásticos que inevitavelmente iriam contribuir para uma maior produção de resíduos.”

Tabela de Dados		
Supermercados com sacos de plástico pagos (579 clientes observados)		
Saco comprado	51% (CI*: 47%-55%)	
Saco reutilizado	49% (CI: 45%-53%)	
Supermercados com sacos de plástico gratuitos (449 clientes observados)		
Saco oferecido	95% (CI: 93%-97%)	
Saco reutilizado	5% (CI: 3%-7%)	
Optimização de sacos		
	Sacos pagos	Sacos gratuitos
Cheio (2 a 3 terços da capacidade)	52% (CI: 46%-58%)	17% (CI: 14%-21%)
Meio (1 a 2 terços da capacidade)	40% (CI: 34%-45%)	60% (CI: 55%-65%)
Quase vazio (até 1 terço da capacidade)	8,5% (CI: 5%-12%)	23% (CI: 19%-27%)

- 6.1** Efetue os cálculos que levaram à apresentação do segundo intervalo.
- 6.2** Admita agora que quer apresentar uma estimativa pontual para a percentagem de sacos reutilizados em supermercados com sacos de plástico pagos com um erro de estimação não superior a 3%. Quantos clientes deve observar se pretender dar uma resposta com 95% de confiança.
- 6.3** Podemos dizer que em supermercados com sacos de plástico pagos, mais de 48% dos clientes reutiliza os sacos? Responda a esta questão com base nos resultados obtidos na alínea anterior e adicionalmente recorrendo a um teste de hipóteses com uma significância de 2,5%.
- 6.4** Resolva o teste de hipóteses da alínea anterior recorrendo ao SPSS.
- 6.5** Calcule um intervalo de confiança (95%) que lhe permita concluir que a utilização de sacos novos é superior nos supermercados com sacos de plástico gratuitos quando comparados com os supermercados com sacos de plástico pagos.
- 7.** O ficheiro ficha4.sav contém os pesos de um conjunto de 40 lincas que foram observados durante um ano. Os lincas foram selecionados aleatoriamente da população de lincas da zona de observação. A amostra é constituída por 13 fêmeas e 37 machos. Averigue se a proporção de machos na população é maior que $\frac{1}{2}$. Responda à questão:
- 7.1** Efetuando um teste de hipóteses com base na região crítica e considerando um nível de significância de 10%;
- 7.2** Utilizando um intervalo de confiança a 90%.
- 8.** Numa região afetada por um surto epidémico, observou-se uma amostra de 2500 indivíduos, tendo-se encontrado 625 contaminados. Teste, ao nível de significância 10%, se a proporção de indivíduos contaminados é diferente de 0.2. Responda à questão:
- 8.1** Efetuando um teste de hipóteses com base na região crítica;
- 8.2** Utilizando um intervalo de confiança a 99%.
- 8.3** Qual deve ser o tamanho da amostra a recolher para que, com 99% de confiança, o erro máximo cometido seja inferior a 0.01.

9. Considere novamente a base de dado compra_calçado_desportivo.sav. Pretende-se investigar se quem tem mais de um par de calçado desportivo de marca tende a gastar mais na compra de um par, para tal admita que o gasto em calçado desportivo tem distribuição normal nas populações em estudo.
- 9.1 Ao nível de significância de 5% teste a igualdade das variâncias recorrendo:
- Ao teste F para o quociente de duas variâncias.
 - Ao SPSS e ao teste de Levene.
- 9.2 Recorrendo ao SPSS responda à questão colocada no enunciado.
- 9.3 Forneça um intervalo de confiança a 95% para a diferença das médias entre os grupos.
10. Testaram-se dois tipos, A e B, de soluções químicas em relação ao ph (grau de acidez da solução). A análise de 40 amostras da solução A acusou ph médio de 7.52 com desvio padrão de 0.024, enquanto a análise de 45 amostras da solução B acusou ph médio de 7.49 com desvio padrão de 0.032. Ao nível de significância de 0.05, teste a hipótese dos dois tipos de solução terem ph diferente.
11. Um biólogo pretende verificar o efeito de um pesticida sobre as larvas de um mosquito. Para o efeito, foram tratadas com pesticida algumas áreas agrícolas e posteriormente contabilizado o número de larvas existentes em charcos de áreas tratadas e não tratadas. O ficheiro ficha4.sav contém os valores obtidos, em número de larvas por charco, divididos por 100. Utilizando o SPSS:
- 11.1 Valide os pressupostos inerentes à realização de um teste t para comparação de médias (considere $\alpha = 0.01$ nos testes de ajustamento que efetuar).
- 11.2 Efetue um teste para averiguar se o número médio de larvas é significativamente diferente nas áreas tratadas e não tratadas ($\alpha = 0.01$).
- Escreva as hipóteses;
 - Registe o valor do p-value do teste t;
 - Conclua.
- 11.3 Efetue um teste t para averiguar se o número médio de larvas é significativamente inferior nas áreas tratadas relativamente às não tratadas ($\alpha = 0.01$).
- Escreva as hipóteses;
 - Determine o valor do p-value do teste t unilateral;
 - Conclua.
- 11.4 Responda à alínea b) utilizando o intervalo de confiança a 99% para a diferença de médias.
12. O ficheiro ficha4.sav contém medições da temperatura (em graus C) e do ritmo cardíaco (em batimentos/minuto) efetuados em 65 homens e 65 mulheres selecionados ao acaso. Considere que a temperatura tem uma distribuição normal nas duas populações. Utilizando o SPSS obteve-se o seguinte output dum teste t de comparação de médias relativamente à variável temperatura:

Group Statistics

	Género	N	Mean	Std. Deviation	Std. Error Mean
Temperatura	m	65	36,72479	,388198	,048150
	f	65	36,88547	,413049	,051232

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Temperatura	Equal variances assumed	,061	,805	-2,285	128	,024	-,160684	,070308	-,299800	-,021568
	Equal variances not assumed			-2,285	127,510	,024	-,160684	,070308	-,299805	-,021563

12.1 Indique o significado de cada um dos campos das tabelas anteriores.

12.2 O que pode dizer sobre a homogeneidade das variâncias tratadas ($\alpha = 0.05$)?

12.3 Considerando $\alpha = 0.05$ e com base na análise do p-value:

- I. averigue se existem diferenças significativas entre as temperaturas de homens e mulheres;
- II. averigue se as temperaturas das mulheres são significativamente superiores às dos homens;
- III. averigue se as temperaturas das mulheres são significativamente inferiores às dos homens;

12.4 Forneça um intervalo de confiança a 95% para a diferença de temperatura média entre os dois sexos.

12.5 Com base no intervalo da alínea anterior confirme a resposta dada na alínea c) i.

13. O ficheiro ficha4.sav contém dados referentes a um ensaio clínico que pretendia estudar a influência do consumo de água e chá nos níveis de colesterol no sangue. Foram selecionados aleatoriamente 37 indivíduos aos quais se mediu o nível de colesterol no início da experiência. Seguidamente os indivíduos tomaram um litro de água por dia e ao fim de um mês voltou-se a medir o colesterol. Finalmente durante mais um mês os mesmos indivíduos tomaram um litro de chá por dia e no final mediu-se novamente o colesterol. Utilizando o SPSS:

13.1 Efetue um teste t para averiguar se o nível de colesterol após o consumo de água é significativamente diferente do nível após o consumo de chá ($\alpha = 0.05$).

- I. Escreva as hipóteses;
- II. Registe o valor do p-value;
- III. Conclua.

13.2 Efetue um teste t para averiguar se o nível de colesterol após o consumo de água é significativamente inferior ao nível após o consumo de chá ($\alpha = 0.05$).

- I. Escreva as hipóteses;
- II. Registe o valor do p-value;
- III. Conclua.

14. Para testar a influência das estações do ano sobre a massa corporal de uma espécie de lince, pesaram-se 40 lince em dois momentos: na estação seca e na estação chuvosa do ano de 2002. Os dados obtidos encontram-se no ficheiro ficha4.sav. Utilizando o SPSS obteve-se o seguinte output relativo a um teste t de comparação de médias em amostras emparelhadas. Considere que foram validadas as condições de aplicabilidade deste teste.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Peso na Estação Chuvosa	17,71	40	2,461	,389
	Peso na Estação Seca	18,94	40	1,797	,284

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	Peso na Estação Chuvosa - Peso na Estação Seca	-1,233	3,351	,530	-2,305	-,162	-2,328	39	,025

Considerando $\alpha = 0.05$ e com base na análise do p-value:

- 14.1 averigue se existem diferenças significativas entre os pesos dos lince nas duas estações;
- 14.2 averigue se os lince pesam mais na estação seca do que na molhada;
- 14.3 averigue se os lince pesam menos na estação seca do que na molhada.

15. O peso de 5 grãos foi medido a partir de duas variedades experimentais designadas “premier” e “super”. Cada grão foi medido em mg com uma casa decimal. Pretende-se averiguar se existem diferenças significativas entre os pesos das duas variedades ($\alpha = 0.05$).

premier	super
24.5	26.4
23.4	27.0
25.3	25.2
23.4	25.8
22.1	27.1

- 15.1 Assumindo que as observações são provenientes de populações normais, efetue um teste t de comparação de médias para responder à questão.
- i. Escreva as hipóteses;
 - ii. Registe o valor do p-value;
 - iii. Conclua.

16. Uma amostra de 29 alturas de plantas de membros de uma determinada espécie forneceu uma variância amostral de 14.62 cm^2 e uma amostra de 25 alturas de plantas de membros de uma outra espécie forneceu uma variância amostral de 8.45 cm^2 .

16.1 Teste, ao nível de significância 5%, se a variância da segunda amostra é significativamente inferior à da primeira.

16.2 Teste, ao nível de significância 10%, e com base no intervalo de confiança se as variâncias das amostras são significativamente diferentes.

17. O diretor de um grupo empresarial pretende comparar as vendas de duas lojas, A e B, do grupo que se dedicam à comercialização do mesmo produto. O diretor acredita que as vendas da loja A são superiores às da loja B. O quadro abaixo apresenta os resultados semanais (em dezenas de Euros), obtidos durante 11 semanas para loja A e 13 semanas para a loja B, relativamente ao número de vendas.

Loja A	123	134	146	104	119	124	161	107	83	113	129		
Loja B	70	118	101	85	107	132	94	100	124	158	95	65	80

O que se pode concluir acerca da suposição do diretor comercial, ao nível de significância de 0.05?

18. Numa determinada empresa, pretende-se controlar o tempo médio de resposta aos pedidos dos seus clientes e também a variabilidade desse tempo. Sabe-se que os tempos de resposta seguem uma distribuição normal. Para uma amostra aleatória de 24 serviços executados recentemente, registaram-se os tempos, tendo-se obtido os seguintes resultados:

Tempos em horas (espaço de tempo entre o registo do pedido do cliente e a finalização do serviço)	
Média	Desvio padrão
9 h	2.7 h

Considere um nível de significância de 1%.

18.1 Considere os conjuntos de hipóteses A, B e C, indicadas na tabela abaixo.

A	B	C
$H_0: \sigma=2.5$	$H_0: \sigma=3$	$H_0: \sigma=2$
$H_1: \sigma>2.5$	$H_1: \sigma>3$	$H_1: \sigma<2$

Perante os resultados obtidos na amostra recolhida, que conjunto de hipóteses faz sentido testar?

Apenas A

Apenas B

Apenas C

Nenhuma

18.2 Considere o teste das hipóteses $H_0: \sigma^2= 25$ vs $H_1: \sigma^2< 25$. Apresente: Valor observado da Estatística de teste e a região crítica

18.3 A hipótese de normalidade explicitada no enunciado (sublinhada) é necessária para o teste de hipóteses da questão anterior?

Sim

Não

18.4 No teste das hipóteses da questão anterior ($H_0: \sigma^2 = 25$ vs $H_1: \sigma^2 < 25$), admita que $p\text{-value} \approx 0$. Complete a seguinte afirmação:

“A amostra recolhida _____ (dá / não dá) evidência de que _____ (a média/ a variância) do tempo de resposta é _____ (superior/inferior) a _____ ($5/25/2.7/2.7^2/9$). “

18.5 Determine um intervalo de confiança a 95% para o tempo médio de resposta aos pedidos dos clientes.

18.6 Quando estimamos o tempo médio de resposta aos pedidos dos clientes recorrendo à média amostral (9 horas) estou a cometer um erro não superior a _____ com 95% de confiança.