41. $\hat{p} = 0{,}52$ e $z = 1{,}960$

Se a amplitude é 0,2, então a margem de erro é 0,1.

$$1{,}960 \times \sqrt{\frac{0{,}5 \times 0{,}48}{n}} = 0{,}1 \Leftrightarrow \sqrt{\frac{0{,}2496}{n}} = 0{,}0510 \Leftrightarrow \frac{0{,}2496}{n} = 0{,}0510^2 \Leftrightarrow$$

$$\Leftrightarrow n = \frac{0{,}2496}{0{,}0510^2} \Leftrightarrow n \approx 96$$

Tema 5 – Exercícios globais (pág. 271)

1.1 $4 + 1 + 1 = 6\%$

Como a dimensão da amostra é 15 800, o número de inquiridos correspondente é $0{,}06 \times 15\ 800 = 948$.

1.2 Calculamos a percentagem acumulada:

Escala	Percentagem	Percentagem acumulada
1	10	10
2	12	22
3	16	38
4	17	55
5	19	74
6	12	86
7	8	94
8	4	98
9	1	99
10	1	100

O primeiro quartil é o valor da variável abaixo do qual se encontram 25% dos dados. Verificamos que corresponde ao nível 3.

A mediana corresponde ao nível 4, pois é o valor da variável abaixo do qual se encontram 50% dos dados.

1.3 $n = 15\ 800$ e $\hat{p} = 0{,}1$

O intervalo de confiança de 99% para a proporção é dado por:

$$\left]\hat{p} - z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}};\ \hat{p} + z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right[= \left]0{,}1 - 2{,}576\sqrt{\frac{0{,}1 \times 0{,}9}{15800}};\ 0{,}1 + 2{,}576\sqrt{\frac{0{,}1 \times 0{,}9}{15800}}\right[=$$

$$= \left]0{,}094;\ 0{,}106\right[$$

1.4 A margem de erro de um intervalo de confiança é metade da sua amplitude.

Considerando $\hat{p} = 0{,}5$ e $n = 100$, o intervalo de confiança de 95% para μ é:

$$\left]0,5-1,96\sqrt{\frac{0,5\times 0,5}{100}};\ 0,5+1,96\sqrt{\frac{0,5\times 0,5}{100}}\right[=]0,402;\ 0,598[$$

A margem de erro neste intervalo de confiança é:

$$\frac{0{,}598-0{,}402}{2} = 0{,}098$$

Se $n = 200$:

$$\left]0,5-1,96\sqrt{\frac{0,5\times 0,5}{200}};\ 0,5+1,96\sqrt{\frac{0,5\times 0,5}{200}}\right[=]0,431;\ 0,569[$$

Se $n = 500$:

$$\left]0,5-1,96\sqrt{\frac{0,5\times 0,5}{500}};\ 0,5+1,96\sqrt{\frac{0,5\times 0,5}{500}}\right[=]0,456;\ 0,544[$$

À medida que a dimensão da amostra aumenta, a margem de erro diminui.

2.1 Sabemos que existe «empate técnico» quando a diferença entre as estimativas pontuais é, em valor absoluto, inferior à margem de erro.

Então, $|41-39| = 2\%$ é menor que a margem de erro, pelo que podemos dizer que estavam em situação de «empate técnico».

2.2 Não. As estimativas das percentagens de votos para os partidos X e Y eram de 39% e 41%, respetivamente, com uma margem de erro de 6% e um nível de confiança de 95%. Assim, seria de esperar que, com 95% de confiança, o partido X tivesse entre 33% e 45% e o partido Y, entre 35% e 47%. A percentagem dos votos do partido X aproximou-se de 45% e a do partido Y de 35%.

2.3 Sabemos que: $\hat{p} = 39\%$.

A margem de erro é:

$$1{,}96\sqrt{\frac{0{,}39\times 0{,}61}{n}} = 0{,}06 \Leftrightarrow \sqrt{\frac{0{,}39\times 0{,}61}{n}} = \frac{0{,}06}{1{,}96} \Leftrightarrow \frac{0{,}2379}{n} = \left(\frac{0{,}06}{1{,}96}\right)^2 \Leftrightarrow n \approx 253{,}87$$

Se a margem de erro passar para 0,03, mantendo o nível de confiança:

$$1{,}96\sqrt{\frac{0{,}39\times 0{,}61}{n}} = 0{,}03 \Leftrightarrow \sqrt{\frac{0{,}39\times 0{,}61}{n}} = \frac{0{,}03}{1{,}96} \Leftrightarrow \frac{0{,}39\times 0{,}61}{n} = \left(\frac{0{,}03}{1{,}96}\right)^2 \Leftrightarrow n \approx 1015$$

A afirmação é falsa, porque a dimensão da amostra quadruplicou.

3.1 Ao restringir-se a amostra aos presentes na sala, estamos a excluir os que tentaram entrar no *site* e os que desistiram. Assim, os resultados provenientes desta amostra poderão ser enviesados, pois os que estavam interessados em aceder ao *site* não fizeram parte da amostra, afetando o valor da percentagem de entrada à primeira tentativa.

Para a amostra ser representativa, esta deveria conter os que tentaram aceder ao *site*, quer tenham ou não conseguido entrar.

3.2 $\hat{p} = \dfrac{39}{50} = 0{,}78$ e $n = 50$

O intervalo de confiança de 95% para a proporção é:

$$\left]0,78 - 1,96\sqrt{\frac{0,78 \times 0,22}{50}};\ 0,78 + 1,96\sqrt{\frac{0,78 \times 0,22}{50}}\right[= \left]0,665;\ 0,895\right[$$

3.3 Introduzindo os valores na calculadora e determinando a regressão linear, obtemos os valores: $a = 3{,}85$ e $b = 4{,}94$.

4.1 A variável em estudo é o comprimento de cada parafuso medido em centímetros.

4.2 O número total de parafusos é 100.

O número de parafusos cujo comprimento é inferior a 5,5 centímetros é $3 + 5 + 13 + 18 = 48$, que corresponde a 48%.

4.3 Para calcular a média, temos de determinar a marca de cada classe:

$$\overline{x} = \frac{5,05 \times 3 + 5,15 \times 5 + 5,25 \times 9 + 5,35 \times 13 + 5,45 \times 18 + 5,55 \times 19 +}{}$$

$$\frac{+5,65 \times 17 + 5,75 \times 10 + 5,85 \times 3 + 5,95 \times 2 + 6,05}{100} \approx 5,5$$

4.4 O menor valor registado é 5,025 e o maior é 6,070, pelo que a amplitude será:

$6{,}070 - 5{,}025 = 1{,}045$

Como pretendemos a existência de sete classes, a amplitude de cada classe será, aproximadamente:

$\dfrac{1{,}045}{7} \approx 0{,}15$

Podemos, então, definir as seguintes classes:

$\left[5,025;\ 5,175\right[;\ \left[5,175;\ 5,325\right[;\ \left[5,325;\ 5,475\right[;\ \left[5,475;\ 5,625\right[;$

$\left[5,625;\ 5,775\right[;\ \left[5,775;\ 5,925\right[;\ \left[5,925;\ 6,075\right[$

Não se tem acesso à distribuição inicial dos dados, sendo que não se consegue distribuir os parafusos pelas novas classes. O estudo inicial foi efetuado com uma distribuição em 11 classes de amplitude 0,1.

4.5 $\overline{x} = 5{,}5$; $s = \sqrt{0{,}043} = 0{,}207$ e $n = \sqrt{100} = 10$

O intervalo de confiança de 95% para a proporção é:

$$\left]5,5 - 1,96 \times \frac{0,207}{10};\ 5,5 + 1,96 \times \frac{0,207}{10}\right[= \left]5,46;\ 5,54\right[$$

5.1 Vamos construir a tabela referente ao sexo masculino com a frequência relativa em percentagem:

Intensidade do gosto de ler	Frequência relativa (%)
Não gosto nada de ler	12
Gosto pouco de ler	26
Gosto de ler de vez em quando	44
Gosto muito de ler	15
Sou viciado na leitura	3

Na tabela referente ao sexo masculino, a moda é «Gosto de ler de vez em quando». No gráfico referente ao sexo feminino, a moda também é «Gosto de ler de vez em quando», com 49%. Assim, a moda é a mesma em ambos os sexos.

Vamos calcular a percentagem dos que revelaram pelo menos algum gosto pela leitura em cada um dos sexos.

$11+49+31+6=97\%$ nas raparigas

$26+44+15+3=88\%$ nos rapazes

Pode concluir-se que as raparigas revelaram um maior gosto pela leitura do que os rapazes.

Então, podemos concluir que a afirmação é verdadeira.

5.2 O intervalo de confiança de 95% para a proporção de estudantes do Ensino Secundário, do continente, que se identificam como sendo apaixonados pela leitura, tendo em conta que $\hat{p}=\frac{221}{4713}=0{,}0469$ e $n=4713$, é:

$$\left]0,0469-1,96\times\sqrt{\frac{0,0469(1-0,0469)}{4713}};\ 0,0469+1,96\times\sqrt{\frac{0,0469(1-0,0469)}{4713}}\right[=$$

$$=\left]0,041;\ 0,053\right[$$

6.1 Introduzindo os valores das listas correspondentes ao rendimento mensal e às despesas com a alimentação na calculadora, obtém-se o valor do coeficiente de correlação linear: $r\approx 0{,}9$.

Atendendo a este valor, a associação linear é positiva forte, pois *r* é positivo e próximo de 1.

6.2.1 Determinando a regressão linear na calculadora, obtemos $a=0{,}1656$ e $b=185{,}1833$.

6.2.2 Usando a reta de regressão linear encontrada na alínea anterior, verificamos que $x=1750$ e $y=474{,}96$.

Podemos estimar que o valor das despesas de alimentação de um agregado familiar, cujo rendimento mensal é 1750 €, é aproximadamente 475 €.

6.3 Para o rendimento mensal, obteve-se a média amostral de 1712,5 e a mediana 1575.

Alterando o valor de 2800 para 8000, a média amostral passou para 2145,83 e a mediana manteve-se.

A média é sensível à alteração de qualquer um dos dados. A mediana permaneceu inalterada mesmo após a diferença introduzida. A mediana é mais resistente.

6.4 $\bar{x} = 270$, $s = 100$ e $n = 50$

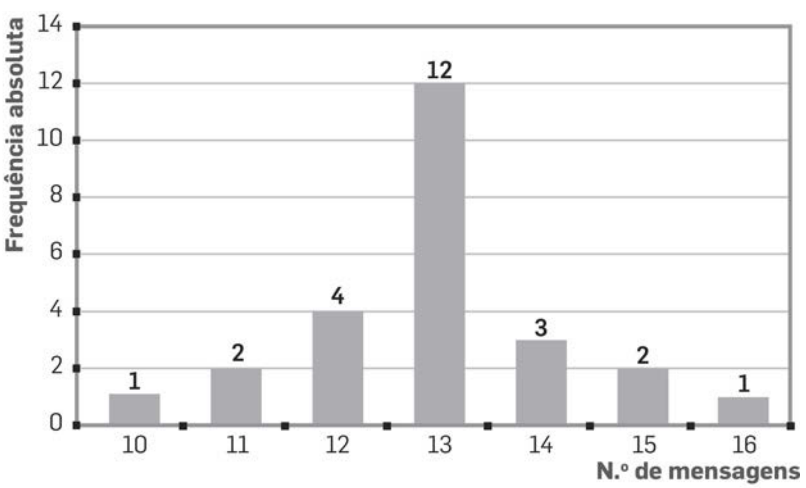
O intervalo de confiança de 95% para o valor médio das despesas com a alimentação é dado por:

$$\left]270 - 1,96 \times \frac{100}{\sqrt{50}};\ 270 + 1,96 \times \frac{100}{\sqrt{50}}\right[= \left]242,28;\ 297,72\right[$$

7.1.1 Tabela referente às frequências relativas simples e às frequências relativas acumuladas do número de mensagens recebidas:

N.º de mensagens	Freq. relativa simples	Freq. relativa acumulada
10	0,04	0,04
11	0,08	0,12
12	0,16	0,28
13	0,48	0,76
14	0,12	0,88
15	0,08	0,96
16	0,04	1
Total	**1**	

7.1.2 Diagrama de barras com a frequência absoluta do número de mensagens recebidas:



7.2 Recorrendo à calculadora:

Turma A: $\bar{x} = 12,96$ e $\sigma = 3,39$

Turma B: $\bar{x} = 12,96$ e $\sigma = 1,28$

Sendo as médias iguais, os desvios-padrão são diferentes porque na turma B o 13 é o que tem maior frequência, sendo que as inferiores ou superiores a 13 têm frequências absolutas menores. Portanto, há uma baixa variabilidade em relação à média.

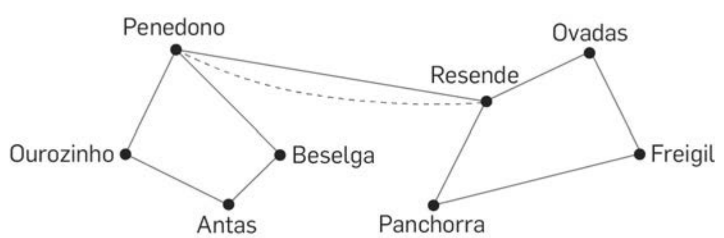
Na turma A, apesar de 13 ser a maior frequência absoluta, existe uma maior amplitude amostral $(19 - 6 = 13)$ do que na turma B $(16 - 10 = 6)$. Portanto, há maior variabilidade dos dados, relativamente à média, na turma A do que na turma B, o que se traduz num desvio-padrão maior.

7.3 $\hat{p}=\frac{125}{250}=0,5$ e $n=250$

O intervalo de confiança de 95% para a proporção de mensagens com extensão de 30 caracteres recebidas no telemóvel pelos alunos da escola é:

$$\left]0,5-1,96\times\sqrt{\frac{0,5(1-0,5)}{250}};\ 0,5+1,96\times\sqrt{\frac{0,5(1-0,5)}{250}}\right[=\left]0,44;0,56\right[$$

8.1 A afirmação é verdadeira, uma vez que no grafo que representa a situação existem vértices com grau ímpar. Assim, este grafo não admite circuitos de Euler. No entanto, se admitirmos a duplicação da aresta que liga Penedono a Resende, obtemos um novo grafo onde todos os vértices têm grau par, pelo que será possível encontrar circuitos de Euler.



8.2 $n=500$, $\overline{x}=830$ e $s=220$

O intervalo de confiança para o valor médio de uma fatura da empresa Silva-Filhos é:

$$\left]830-2,576\times\frac{220}{\sqrt{500}};\ 830+2,576\times\frac{220}{\sqrt{500}}\right[=\left]804,66;\ 855,34\right[$$

Há razões para duvidar da afirmação do gerente, visto que 800 € para o valor médio de uma fatura da empresa não pertence ao intervalo de confiança de 99%.

9. Sabemos que: $n=5000$ e $\hat{p}=0,41$

O intervalo de confiança de 99% para a proporção de dadores com o grupo sanguíneo O é dado por:

$$\left]0,41-2,576\times\sqrt{\frac{0,41\times0,59}{5000}};\ 0,41+2,576\times\sqrt{\frac{0,41\times0,59}{5000}}\right[=\left]0,3921;\ 0,4279\right[$$

O intervalo requerido é $\left]39,21\%;\ 42,79\%\right[$.

Tema 5 – Teste final (pág. 278)

1.1 População: todos os eleitores votantes em Portugal

Amostra: os 1354 eleitores selecionados em Portugal para a entrevista

1.2 A amostra maior pode não ser a mais representativa da população. Teríamos de conhecer o(s) método(s) utilizado(s) na sua recolha, que influencia(m) a qualidade e representabilidade da população. Poderíamos apenas dizer que, como a amostra de 1354 eleitores foi recolhida em