Introdução à inferência estatística

Pág. 200

1.1. Este tipo de estudos – as sondagens eleitorais – têm como objetivo aferir o sentido de voto dos eleitores. Isto permite, não só uma monitorização da realidade política e social, mas também que as forças políticas e os candidatos a cargos políticos possam avaliar a eficiência das estratégias de campanha que estão a seguir, o seu próprio desempenho político e a popularidade.

1.2. Devido à dimensão da população, a inquirição de todos os elementos que a constituem, para além de demorar muito tempo, também envolveria certamente custos muito avultados. Tudo isto tornaria o processo impraticável. Nesta situação particular, toda a população será inquirida, salvo as abstenções, no dia das eleições, ou seja, em vez de uma sondagem estaríamos perante o próprio ato eleitoral em si. Obviamente que os resultados apresentados na sondagem podem ser diferentes daqueles que na realidade irão produzir-se, sendo que a qualidade da sondagem dependerá da forma como ela foi construída e realizada, assumindo a escolha da amostra um papel fulcral.

1.3. Selecionaria uma amostra aleatória e representativa dos alunos da escola, utilizando os métodos de amostragem estudados no 10.º ano. Por exemplo, seriam selecionados, aleatoriamente e de forma proporcional ao que ocorre na população, alunos de ambos os sexos e de todos os anos de escolaridade. Aos alunos selecionados seria então aplicado o inquérito ou pergunta.

2.1. $\mu = 0{,}25 \times 0 + 0{,}40 \times 1 + 0{,}28 \times 2 + 0{,}05 \times 3 + 0{,}02 \times 4$
$= 1{,}19$

Interpretação: O guarda-redes sofre, em média, $1{,}19$ golo por jogo.

2.2. $P(X \geq 2) = P(X = 2) + P(X = 3) + P(X = 4) =$
$= 0{,}28 + 0{,}05 + 0{,}02 = 0{,}35$

A probabilidade pedida é 0,35 (ou 35%).

2.3. $\sigma^2 = (0 - 1{,}19)^2 \times 0{,}25 + (1 - 1{,}19)^2 \times 0{,}40 +$
$+ (2 - 1{,}19)^2 \times 0{,}28 + (3 - 1{,}19)^2 \times 0{,}05 +$
$+ (4 - 1{,}19)^2 \times 0{,}02$

$\sigma = \sqrt{0{,}2975 + 0{,}0144 + 0{,}1837 + 0{,}1638 + 0{,}1579}$
$= 0{,}9348$

$]1{,}19 - 0{,}9348\ ;\ 1{,}19 + 0{,}9348[\ =\ ]0{,}2552\ ;\ 2{,}1248[$

$P(0{,}2552 < X < 2{,}1248) = P(X = 1) + P(X = 2) =$
$= 0{,}40 + 0{,}28 = 0{,}68$

$0{,}68 \times 50 = 34$

Em 34 jogos.

Repare que, se admitirmos que a distribuição é normal obtemos um resultado igual uma vez que sabemos que:
$P(\mu - \sigma < X < \mu + \sigma) \approx 0{,}6827$.

Consequentemente, $0{,}6827 \times 50 \approx 34$ jogos.

3.1. Admitamos que X representa o *peso* dos alunos do 11.º ano da escola. Sabemos que $P(X \geq 62) = 50\%$.
No entanto, não podemos afirmar com certeza absoluta que metade dos alunos *pesa* menos de 62 kg. Apenas podemos dizer que a probabilidade de selecionar um aluno ao acaso que *pese* pelo menos 62 kg é de 50%, ou seja, que se espera que 45 alunos *pesem* pelo menos 62 kg.

3.2. $P(X < 54) = 0{,}5 - P(54 < X < 62) =$
$= 0{,}5 - \dfrac{P(54 < X < 70)}{2} = 0{,}5 - \dfrac{0{,}6827}{2}$
$= 0{,}158\,65 \approx 15{,}9\%$

3.3. Recorrendo à calculadora: $P(58 < X < 65) \approx 0{,}3376$
$0{,}3376 \times 90 \approx 30$ alunos com *peso* compreendido entre 58 kg e 65 kg.

Proporção: $\dfrac{30}{90} = \dfrac{1}{3}$

Aproximadamente, $\dfrac{1}{3}$ dos alunos.

Pág. 203

2.1. População: todos os alunos da escola.

2.2. Proporção de alunos da escola que se revelam muito satisfeitos com as condições da biblioteca.

2.3. Proporção de alunos da amostra que se revelam muito satisfeitos com as condições da biblioteca.

Valor da estatística: $\dfrac{42}{60} = \dfrac{7}{10} = 0{,}7 = 70\%$

2.4. Podemos utilizar a estatística para prever/estimar que 70% dos alunos da escola revelam-se muito satisfeitos com as condições da biblioteca.

Assim, $0{,}70 \times 450 = 315$.

Prevê-se que 315 alunos da escola se revelem muito satisfeitos com as condições da biblioteca.

Pág. 207

3.1. $5^2 = 25$ amostras de dois elementos.

3.2.

	30	**28**	**24**	**32**	**26**
30	(30, 30)	(30, 28)	(30, 24)	(30, 32)	(30, 26)
28	(28, 30)	(28, 28)	(28, 24)	(28, 32)	(28, 26)
24	(24, 30)	(24, 28)	(24, 24)	(24, 32)	(24, 26)
32	(32, 30)	(32, 28)	(32, 24)	(32, 32)	(32, 26)
26	(26, 30)	(26, 28)	(26, 24)	(26, 32)	(26, 26)

Médias	**30**	**28**	**24**	**32**	**26**
30	30	29	27	31	28
28	29	28	26	30	27
24	27	26	24	28	25
32	31	30	28	32	29
26	28	27	25	29	26

$\bar{X} = x_i$	24	25	26	27	28	29	30	31	32
$P(\bar{X} = x_i)$	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{3}{25}$	$\frac{4}{25}$	$\frac{1}{5}$	$\frac{4}{25}$	$\frac{3}{25}$	$\frac{2}{25}$	$\frac{1}{25}$

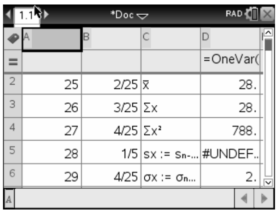
3.3. $\mu = \dfrac{30+28+24+32+26}{5} = 28$

$\sigma = \sqrt{\dfrac{(30-28)^2+(28-28)^2+(24-28)^2+(32-28)^2+(26-28)^2}{5}}$

$= \sqrt{8}$

Recorrendo à calculadora gráfica obtemos:

$\mu_{\overline{X}} = 28$ e $\sigma_{\overline{X}} = 2$



Note-se que: $\sigma_{\overline{X}} = 2 = \sqrt{4} = \dfrac{\sqrt{8}}{\sqrt{2}}$

Conclusão: $\mu_{\overline{X}} = \mu$ e $\sigma_{\overline{X}} = \dfrac{\sigma}{\sqrt{n}}$

Pág. 208

4.1. 200^{40} amostras

4.2. Como $n \geq 30$, o teorema do limite central garante que a distribuição de amostragem da média pode ser aproximada por uma distribuição normal com valor médio μ e desvio-padrão $\dfrac{\sigma}{\sqrt{n}}$. Assim, o valor médio é 165 cm e o desvio-padrão $\dfrac{10}{\sqrt{40}} = \dfrac{\sqrt{10}}{2}$ cm .

Pág. 209

5. Pretendemos obter uma estimativa para o valor médio, logo basta determinar a média da amostra recolhida.

$\overline{x} = \dfrac{2\times0+2\times1+3\times2+4\times3+9\times4+3\times5+5\times6+5\times7+3\times8}{36}$

$\approx 4,4$

Estima-se que são comunicadas, em média, cerca de 4,4 avarias por dia.

Pág. 211

6.1. Tendo em conta o teorema do limite central, espera-se que possa ser modelada por uma distribuição normal com valor médio $\mu = 48$ kg e desvio-padrão $\sigma = \dfrac{12}{\sqrt{100}} = 1,2$ kg .

Podemos escrever que $\overline{X} \sim N(48;\ 1,2)$.

6.2. Pretende-se determinar $P(50 < \overline{X} < 60)$.

Tendo em conta que $\overline{X} \sim N(48;\ 1,2)$, recorrendo à calculadora gráfica, obtém-se $P(50 < \overline{X} < 60) \approx 0,048$.

A probabilidade pedida é, aproximadamente, 4,8% .

Pág. 212

7.1. Admitamos que X representa a mesada, em euros, dos adolescentes. Sabe-se que $X \sim N(16,55;\ 7,50)$.

O teorema do limite central garante que $\overline{X} \sim N\left(16,55;\ \dfrac{7,50}{\sqrt{25}}\right)$, ou seja, $\overline{X} \sim N(16,55;\ 1,50)$.

Recorrendo à calculadora gráfica calculamos:

$P(\overline{X} > 20) = 0,5 - P(16,55 < \overline{X} < 20) \approx 0,5 - 0,489 =$
$= 0,011 = 1,1\%$

A probabilidade pedida é, aproximadamente, 1,1% .

7.2. Pretende-se determinar $P(|\overline{X} - \mu| \leq 0,50)$.

Seja $Z \sim N(0,\ 1)$, então $Z = \dfrac{\overline{X} - 16,55}{1,50} \Leftrightarrow \overline{X} = 1,50Z + 16,55$

Assim:

$P\left(\left|\overline{X} - \mu\right| \leq 0,50\right) = P(|1,50Z + 16,55 - 16,55| \leq 0,50) =$

$= P(|1,50Z| \leq 0,50) = P\left(-\dfrac{0,50}{1,50} \leq Z \leq \dfrac{0,50}{1,50}\right) \approx 0,2611$

A probabilidade pedida é, aproximadamente, 0,2611 .

Pág. 215

8.1. a) $\left]12 - 1,96\times\dfrac{8,5}{\sqrt{150}};\ 12 + 1,96\times\dfrac{8,5}{\sqrt{150}}\right[\approx$

$\approx]10,640;\ 13,360[$

b) $\left]12 - 2,576\times\dfrac{8,5}{\sqrt{150}};\ 12 + 2,576\times\dfrac{8,5}{\sqrt{150}}\right[\approx$

$\approx]10,212;\ 13,788[$

8.2.

Grau de confiança	Amplitude
90%	2,284
95%	2,72
99%	3,576

Quanto maior for o grau de confiança, maior é a amplitude do intervalo.

9. $\left]73,8 - 1,96\times\dfrac{25}{\sqrt{16}};\ 73,8 + 1,96\times\dfrac{25}{\sqrt{16}}\right[=]61,55\ ;\ 86,05[$

Pág. 216

10. $\left]1,70 - 1,96\times\dfrac{0,40}{\sqrt{80}};\ 1,70 + 1,96\times\dfrac{0,40}{\sqrt{80}}\right[\approx]1,61;\ 1,79[$

Tendo em conta o estudo efetuado pelo seu filho, é possível dizer-se que o sr. Arlindo não tem razão uma vez que se pode estimar, com uma confiança de 95% , que o valor médio dos talões está compreendido entre 1,61 € e 1,79 € .

11. Pretende-se determinar o valor de z tal que:

$\left]\overline{x} - z\times\dfrac{120}{\sqrt{500}};\ \overline{x} + z\times\dfrac{120}{\sqrt{500}}\right[=]986,176;\ 1013,82[$

Observe-se que $z\times\dfrac{s}{\sqrt{n}}$ é igual metade da amplitude do intervalo (raio do intervalo). Assim:

$z\times\dfrac{120}{\sqrt{500}} = \dfrac{1013,82 - 986,176}{2} \Leftrightarrow z\times\dfrac{120}{\sqrt{500}} = 13,822 \Leftrightarrow$

$\Leftrightarrow z = \dfrac{13,822\times\sqrt{500}}{120}$

Então $z = 2,576$.

O nível de confiança é 99% .

Pág. 217

12.1. **População**: todos os alunos da E.S. de Beiraliz.

Amostra: os 80 alunos da E.S. de Beiraliz selecionados.

12.2. Vamos usar a proporção amostral, $\hat{p}$, para estimar a proporção populacional, p.

$$\hat{p} = \frac{12+38}{80} = \frac{5}{8} = 0,625$$

Estima-se que $62,5\%$ dos alunos não vão votar na lista A.

Pág. 219

13.1. Proporção de alunos com menos de 10 valores nos exames de Biologia: $p = 0,20$

Uma vez que a dimensão da amostra é no mínimo 30, o teorema do limite central garante que a distribuição de amostragem da proporção pode ser aproximada por um modelo normal de valor médio $p = 0,20$ e desvio-padrão $\sqrt{\dfrac{0,20(1-0,20)}{75}} \approx 0,046$.

Assim, a proporção amostral, $\hat{p}$, segue, aproximadamente, um modelo normal $N(0,20;\ 0,046)$.

13.2. Atendendo a que $\hat{p} \sim N(0,20; 0,046)$:

$P(\hat{p} \leq 0,18) = 0,5 - P(0,18 \leq \hat{p} \leq 0,20) \approx 0,5 - 0,17 = 0,33$

A probabilidade pedida é, aproximadamente, 33%.

Pág. 221

14. Proporção amostral: $\hat{p} = \dfrac{125}{250} = 0,5$

Intervalo de confiança:

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

$\approx\ ]0,44;\ 0,56[$

15.1. $$\left]0,5 - 1,96 \times \sqrt{\frac{0,5(1-0,5)}{400}};\ 0,5 + 1,96 \times \sqrt{\frac{0,5(1-0,5)}{400}}\right[\approx$$

$\approx\ ]0,45;\ 0,55[$

15.2. Amplitude do intervalo do exercício **14.**: $0,56 - 0,44 = 0,12$

Amplitude do intervalo do exercício **15.1.**: $0,55 - 0,45 = 0,10$

Podemos concluir que aumentando a dimensão da amostra, e mantendo o grau de confiança, a amplitude do intervalo diminui (o intervalo de valores é mais "restrito").

Pág. 223

16.1. Intervalo de confiança:

$$\left]48,5 - 1,645 \times \frac{4,3}{\sqrt{70}};\ 48,5 + 1,645 \times \frac{4,3}{\sqrt{70}}\right[\approx\ ]47,7\ ;\ 49,3[$$

Margem de erro: $E = 1,645 \times \dfrac{4,3}{\sqrt{70}} \approx 0,8$

16.2. Pretende-se determinar o valor de n tal que:

$$1,645 \times \frac{4,3}{\sqrt{n}} = 0,5 \Leftrightarrow 1,645 \times 4,3 = 0,5\sqrt{n} \Leftrightarrow \sqrt{n} = \frac{1,645 \times 4,3}{0,5}$$

Assim, $n = \left(\dfrac{1,645 \times 4,3}{0,5}\right)^2 \approx 200,138$.

Alternativamente, pode-se aplicar diretamente a fórmula apresentada no manual $n = \left(\dfrac{z \times \sigma}{E}\right)^2$.

A amostra deverá ter pelo menos 201 funcionários.

Pág. 224

17. Aplicando a fórmula apresentada no manual, obtém-se:

$$n = \left(\frac{2,576}{0,025}\right)^2 \times 0,15 \times (1-0,15) \approx 1353,7$$

Devem ser inquiridas 1354 pessoas.

Pág. 228

18.1. A população em estudo é o conjunto de todos os alunos portugueses do 11.º ano de escolaridade que frequentam a disciplina de Matemática Aplicada às Ciências Sociais.

18.2. Os parâmetros valor médio da CIF na disciplina de MACS e a proporção de alunos portugueses com CIF negativa na disciplina de MACS.

18.3. A média e a proporção amostral.

CIF média na amostra: $13,6$ valores

Proporção de alunos da amostra com CIF negativa: $0,14$ (ou 14%)

18.4. Utiliza-se a estatística proporção amostral para estimar a proporção populacional.

$1 - \hat{p} = 1 - 0,14 = 0,86$

Prevê-se que 86% dos alunos tenham CIF positiva na disciplina de MACS.

19.1. $\mu = \dfrac{35 + 38,5 + 41,1 + 43}{4} = \dfrac{157,6}{4} = 39,4$

$$\sigma = \sqrt{\frac{(35-39,4)^2 + (38,5-39,4)^2 + (41,1-39,4)^2 + (43-39,4)^2}{4}} =$$

$$= \sqrt{9,005} \approx 3$$

Valor médio: $39,4$

Desvio-padrão: aproximadamente, 3

19.2.

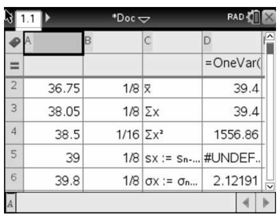
	35	**38,5**	**41,1**	**43**
35	(35; 35)	(35; 38,5)	(35; 41,1)	(35; 43)
38,5	(38,5; 35)	(38,5; 38,5)	(38,5; 41,1)	(38,5; 43)
41,1	(41,1; 35)	(41,1; 38,5)	(41,1; 41,1)	(41,1; 43)
43	(43; 35)	(43; 38,5)	(43; 41,1)	(43; 43)

Médias	**35**	**38,5**	**41,1**	**43**
35	35	36,75	38,05	39
38,5	36,75	38,5	39,8	40,75
41,1	38,05	39,8	41,1	42,05
43	39	40,75	42,05	43

$\bar{X} = x_i$	35	36,75	38,05	38,5	39
$P(\bar{X} = x_i)$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{8}$

$\bar{X} = x_i$	39,8	40,75	41,1	42,05	43
$P(\bar{X} = x_i)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$

19.3. Recorrendo à calculadora gráfica obtemos $\mu_{\bar{X}} = 39,4$ e $\sigma_{\bar{X}} \approx 2,12$.



19.4. $\mu_{\bar{X}} = \mu = 39,4$ e

$$\sigma_{\bar{X}} = 2,12 = \frac{3}{\sqrt{2}} = \frac{\sigma}{\sqrt{n}}$$

19.5. Sim, o estimador é não enviesado (ou cêntrico) pois $\mu_{\bar{X}} = \mu$, ou seja, o valor médio da distribuição de amostragem da média é igual ao valor médio da população.

20. Como $n \geq 30$, o teorema do limite central garante que a distribuição de amostragem da média pode ser aproximada por uma distribuição normal com valor médio μ e desvio-padrão $\frac{\sigma}{\sqrt{n}}$. Assim, o valor médio é 40 cm e o desvio-padrão $\frac{7}{\sqrt{65}} \approx 0,9$ cm, ou seja, $\bar{X} \sim N(40;\ 0,9)$.

A distribuição de amostragem da média pode ser aproximada por uma distribuição normal com valor médio 40 cm e desvio-padrão 0,9 cm.

21. Pretendemos obter uma estimativa para o valor médio, logo basta determinar a média da amostra recolhida.

$$\bar{x} = \frac{12\times0+10\times1+8\times2+6\times3+4\times4}{40} = 1,5$$

Estima-se que o valor médio do número de defeitos nos *microchips* produzidos pela fábrica seja de 1,5 defeito.

22.1. Como $n \geq 30$, o teorema do limite central garante que a distribuição de amostragem da média pode ser aproximada por uma distribuição normal com valor médio μ e desvio-padrão $\frac{\sigma}{\sqrt{n}}$. Assim, o valor médio é 8520 e o desvio-padrão $\frac{180}{\sqrt{36}} = 30$, ou seja, $\bar{X} \sim N(8520,\ 30)$. Espera-se que a distribuição de amostragem da média possa ser aproximada por uma distribuição normal com valor médio 8520 e desvio-padrão 30.

22.2. Atendendo a que $\bar{X} \sim N(8520,\ 30)$ e com o recurso à calculadora gráfica:

$P(\bar{X} > 8490) = 0,5 + P(8490 < \bar{X} < 8520) \approx$

$\approx 0,5 + 0,3413 = 0,8413$

A probabilidade pedida é, aproximadamente, 84,13%.

Pág. 229

23.1. Admitamos que X representa a pontuação obtida nos testes. Sabe-se que $X \sim N(85;\ 8)$, logo o teorema do limite central garante que a distribuição de amostragem da média pode ser aproximada por uma distribuição normal com valor médio μ e desvio-padrão $\frac{\sigma}{\sqrt{n}}$. Assim, $\bar{X} \sim N\left(85;\ \frac{8}{\sqrt{25}}\right)$, ou seja, $\bar{X} \sim N(85;\ 1,6)$. Então:

$P(\bar{X} < 82) = 0,5 - P(82 < \bar{X} < 85) \approx 0,5 - 0,470 = 0,03$

A probabilidade pedida é, aproximadamente, 3%.

23.2. Pretende-se determinar $P(|\bar{X} - \mu|) \leq 5$. Seja $Z \sim N(0,\ 1)$, então $Z = \frac{\bar{X}-85}{1,6} \Leftrightarrow \bar{X} = 1,6Z + 85$. Assim:

$$P(|\bar{X} - \mu| \leq 5) = P(|1,6Z + 85 - 85| \leq 5) = P(|1,6Z| \leq 5)$$
$$= P\left(-\frac{5}{1,6} \leq Z \leq \frac{5}{1,6}\right) \approx 0,9982$$

A probabilidade pedida é, aproximadamente, 0,9982.

24.1. $\left]320 - 1,96\times\frac{35}{\sqrt{20}};\ 320 + 1,96\times\frac{35}{\sqrt{20}}\right[\approx]304,7;\ 335,3[$

24.2. Significa que, com uma confiança de 95%, estima-se que o valor médio do número de amigos que cada aluno do 11.º ano da E.S. de Beiraliz tem na rede social está compreendido entre 304,7 e 335,3.

25.1. Intervalo de confiança de 90%:

$\left]17,2 - 1,645\times\frac{7,5}{\sqrt{40}};\ 17,2 + 1,645\times\frac{7,5}{\sqrt{40}}\right[\approx]15,2;\ 19,2[$

Intervalo de confiança de 99%:

$\left]17,2 - 2,576\times\frac{7,5}{\sqrt{40}};\ 17,2 + 2,576\times\frac{7,5}{\sqrt{40}}\right[\approx]14,1;\ 20,3[$

25.2. Amplitude do intervalo de confiança de 90%:

$19,2 - 15,2 = 4$

Amplitude do intervalo de confiança de 99%:

$20,3 - 14,1 = 6,2$

Ao aumentar o nível de confiança, de 90% para 99%, a amplitude do respetivo intervalo de confiança também aumenta.

26. Pretende-se determinar os valores de $\bar{x}$ e s tal que:

$\left]\bar{x} - 1,96\times\frac{s}{\sqrt{500}};\ \bar{x} + 1,96\times\frac{s}{\sqrt{500}}\right[=]32,5;\ 40,3[$

Observe-se que $1,96\times\frac{s}{\sqrt{500}}$ é igual a metade da amplitude do intervalo (raio do intervalo), ou seja, $\frac{40,3-32,5}{2} = 3,9$.

Assim, $1,96\times\frac{s}{\sqrt{500}} = 3,9 \Leftrightarrow s = \frac{3,9\times\sqrt{500}}{1,96}$.

Então $s \approx 44,5$. O valor de $\bar{x}$ determina-se facilmente pois:

$\bar{x} - 3,9 = 32,5 \Leftrightarrow \bar{x} = 32,5 + 3,9 \Leftrightarrow \bar{x} = 36,4$

Conclusão: $\bar{x} = 36,4$ e $s \approx 44,5$.

Pág. 230

27. Vamos usar a proporção amostral, $\hat{p}$, para estimar a proporção populacional, p.

$\hat{p} = \frac{65}{80} = \frac{13}{16} = 0,8125$ e $1 - \hat{p} = 1 - 0,8125 = 0,1875$

Estima-se que 18,75% dos alunos não visitaram a exposição.

28.1. Proporção de munícipes que consideram aceitável a qualidade dos serviços de recolha do lixo: $p = 0,48$

Uma vez que a dimensão da amostra é no mínimo 30, o teorema do limite central garante que a distribuição de amostragem da proporção pode ser aproximada por um modelo normal de valor médio $p = 0,48$ e desvio-padrão $\sqrt{\frac{0,48(1-0,48)}{250}} \approx 0,03$.

Assim, a proporção amostral, $\hat{p}$, segue aproximadamente um modelo normal $N(0,48;\ 0,03)$.

28.2. Atendendo a que $\hat{p} \sim N(0,48;\ 0,03)$:

$P(\hat{p} \leq 0,45) = 0,5 - P(0,45 \leq \hat{p} \leq 0,48) \approx 0,5 - 0,341 =$

$= 0,159$

A probabilidade pedida é, aproximadamente, 15,9%.

29.1. Seja p a proporção de trabalhadores do sexo masculino.

Sabe-se que $p = 1 - \frac{1}{4} = 0,75$.

Uma vez que a dimensão da amostra é no mínimo 30 , o teorema do limite central garante que a distribuição de amostragem da proporção pode ser aproximada por um modelo normal de valor médio $p = 0,75$ e desvio-padrão $\sqrt{\frac{0,75(1-0,75)}{40}} \approx 0,0685$.

Assim, a proporção amostral, $\hat{p}$, segue, aproximadamente, um modelo normal $N(0,75;\ 0,0685)$.

$P(\hat{p} < 0,65) = 0,5 - P(0,65 \leq \hat{p} \leq 0,75) \approx 0,5 - 0,4278 =$
$= 0,0722$

Então, $0,0722 \times 300 \approx 22$ amostras.

Espera-se encontrar, aproximadamente, 22 amostras.

29.2. Seja p a proporção de trabalhadores do sexo feminino.

Sabe-se que $p = 0,25$.

Uma vez que a dimensão da amostra é no mínimo 30 , o teorema do limite central garante que a distribuição de amostragem da proporção pode ser aproximada por um modelo normal de valor médio $p = 0,25$ e desvio-padrão $\sqrt{\frac{0,25(1-0,25)}{60}} \approx 0,0685$.

Assim, a proporção amostral, $\hat{p}$, segue aproximadamente um modelo normal $N(0,25;\ 0,0685)$.

$P(0,20 < \hat{p} < 0,30) \approx 0,5346$

Então, $0,5346 \times 300 \approx 160$ amostras.

Espera-se encontrar, aproximadamente, 160 amostras.

30. Proporção amostral: $\hat{p} = \frac{18}{30} = 0,6$

Intervalo de confiança:

$$\left]0,6 - 1,645 \times \sqrt{\frac{0,6(1-0,6)}{30}};\ 0,6 + 1,645 \times \sqrt{\frac{0,6(1-0,6)}{30}}\right[\approx$$

$\approx \left]0,45;\ 0,75\right[$

Com 90% de confiança, estima-se que a proporção de alunos do sexo feminino na escola esteja entre 45% e 75% , aproximadamente.

31.1. $$\left]0,6 - 1,645 \times \sqrt{\frac{0,6(1-0,6)}{100}};\ 0,6 + 1,645 \times \sqrt{\frac{0,6(1-0,6)}{100}}\right[\approx$$

$\approx \left]0,52;\ 0,68\right[$

31.2. Podemos concluir que ao aumentar o tamanho da amostra, mantendo o nível de confiança, a amplitude do intervalo de confiança diminui (estimativa mais precisa).

Pág. 231

32.1. $$\left]12,50 - 1,96 \times \frac{3,5}{\sqrt{50}};\ 12,50 + 1,96 \times \frac{3,5}{\sqrt{50}}\right[\approx \left]11,53;\ 13,47\right[$$

Com 95% de confiança, estima-se que os clientes da loja gastem, em média, entre 11,53 € e 13,47 € , aproximadamente.

Margem de erro: $E = 1,96 \times \frac{3,5}{\sqrt{50}} \approx 0,97$

32.2. Pretende-se determinar o valor de n tal que:

$$2,576 \times \frac{3,5}{\sqrt{n}} = 1,5 \Leftrightarrow 2,576 \times 3,5 = 1,5\sqrt{n} \Leftrightarrow \sqrt{n} = \frac{2,576 \times 3,5}{1,5}$$

Assim, $n = \left(\frac{2,576 \times 3,5}{1,5}\right)^2 \approx 36,1$.

Alternativamente, pode-se aplicar diretamente a fórmula apresentada no manual $n = \left(\frac{z \times \sigma}{E}\right)^2$.

A amostra deverá ter dimensão 37 .

33.1. $\hat{p} = 0,16 + 0,22 + 0,13 = 0,51$

$$\left]0,51 - 1,96 \times \sqrt{\frac{0,51(1-0,51)}{200}};\ 0,51 + 1,96 \times \sqrt{\frac{0,51(1-0,51)}{200}}\right[$$

$\approx \left]0,44;\ 0,58\right[$

33.2. Sabe-se que $\hat{p} = 0,15$.

$$\left]0,15 - 2,576 \times \sqrt{\frac{0,15(1-0,15)}{200}};\ 0,15 + 2,576 \times \sqrt{\frac{0,15(1-0,15)}{200}}\right[$$

$\approx \left]0,08;\ 0,22\right[$

33.3. Aplicando a fórmula apresentada no manual, obtém-se:

$$n = \left(\frac{1,96}{0,01}\right)^2 \times 0,51 \times (1-0,51) \approx 9600,16$$

Terá que se lançar o dado 9601 vezes.

34.1. Se selecionarmos aleatoriamente várias amostras com a mesma dimensão e construirmos os intervalos de confiança respetivos, cerca de 99% desses intervalos irão conter o parâmetro a estimar (o restante 1% não o conterá).

O intervalo $]0,28;\ 0,40[$ é um intervalo de confiança de 99% para a proporção de empresários que considera que a estabilidade do emprego no setor se manteve no último ano.

O que se espera é que este intervalo seja um dos 99% que contêm a proporção que se pretende estimar.

34.2. $E = \frac{0,40 - 0,28}{2} = 0,06$

A margem de erro é de seis pontos percentuais.

35.1. Se a dimensão da amostra aumenta, a margem de erro diminui (para um mesmo nível de confiança).

35.2. Se o nível de confiança aumenta, a margem de erro também aumenta (mantendo a dimensão da amostra).

35.3. Se o desvio-padrão aumenta, a margem de erro também aumenta.

Pág. 232

1.1. **População**: todos os alunos da turma 11.º X.

Amostra: os oito alunos da turma 11.º X selecionados aleatoriamente.

1.2. a) Podemos utilizar a média amostral, $\hat{x}$, como estimativa pontual do valor médio, μ .

$$\overline{x} = \frac{3,50 + 2,30 + 5,15 + 1,90 + 7,10 + 4,45 + 3,20 + 2,90}{8}$$

$\approx 3,81$ €

Estima-se que os alunos da turma 11.º X trazem na carteira, em média, 3,81 € .

b) Podemos utilizar a proporção amostral, $\hat{p}$, como estimativa pontual da proporção populacional, p.

$\hat{p} = \frac{2}{8} = \frac{1}{4} = 0,25$

Estima-se que 25% dos alunos da turma 11.º X têm olhos verdes.

2.1. $E = \frac{41,741 - 39,859}{2} = 0,941$

A margem de erro é 0,941 .

2.2. $z \times \frac{12}{\sqrt{625}} = 0,941 \Leftrightarrow z = \frac{0,941 \times \sqrt{625}}{12}$

Então $z \approx 1,96$.

Nível de confiança do intervalo: 95%

2.3. $\bar{x} - 0,941 = 39,859 \Leftrightarrow \bar{x} = 39,859 + 0,941 \Leftrightarrow \bar{x} = 40,8$

Média amostral: 40,8

3.1. Como $n \geq 30$, o teorema do limite central garante que a distribuição de amostragem da média pode ser aproximada por uma distribuição normal com valor médio μ e desvio-padrão $\frac{\sigma}{\sqrt{n}}$.

Assim, o desvio-padrão é $\frac{6,2}{\sqrt{60}} \approx 0,8$.

3.2. $\left]172,7 - 1,96 \times \frac{6,2}{\sqrt{60}};\ 172,7 + 1,96 \times \frac{6,2}{\sqrt{60}}\right[\approx \left]171,1;\ 174,3\right[$

3.3. O fornecedor deve estar satisfeito uma vez que sabe, com 95% de confiança, que o valor médio (a altura média da população) está compreendido entre 171,1 cm e 174,3 cm (e os moldes estão feitos para uma altura média de 172 cm).

Pág. 233

4.1. $(1 - 0,06) \times 16\,000 = 15\,040$

Espera-se que 15 040 litros de vinho sejam de primeira categoria.

4.2. Uma vez que a dimensão da amostra é no mínimo 30 , o teorema do limite central garante que a distribuição de amostragem da proporção pode ser aproximada por um modelo normal de valor médio $p = 0,06$ e desvio-padrão $\sqrt{\frac{0,06(1-0,06)}{16\,000}} \approx 0,0019$.

Assim, a proporção amostral, $\hat{p}$, segue aproximadamente um modelo normal $N(0,06;\ 0,0019)$.

4.3. Atendendo a que $\hat{p} \sim N(0,06;\ 0,0019)$:

$P\left(\frac{950}{16\,000} < \hat{p} < \frac{1100}{16\,000}\right) \approx 0,6289$

A probabilidade pedida é, aproximadamente, 62,9% .

5.1. Sabe-se que $\hat{p} = 0,30$.

$\left]0,30 - 1,645 \times \sqrt{\frac{0,30(1-0,30)}{3200}};\ 0,30 + 1,645 \times \sqrt{\frac{0,30(1-0,30)}{3200}}\right[$

$\approx \left]0,29;\ 0,31\right[$

5.2. Sabe-se que $\bar{x} = 2,6$.

$\left]2,6 - 1,96 \times \frac{0,5}{\sqrt{3200}};\ 2,6 + 1,96 \times \frac{0,5}{\sqrt{3200}}\right[\approx \left]2,58;\ 2,62\right[$

5.3. $E = 1,96 \times \frac{0,5}{\sqrt{3200}} \approx 0,02$

Margem de erro: aproximadamente, 0,02 .

5.4. Aplicando a fórmula apresentada no manual, obtém-se:

$n = \left(\frac{2,576 \times 0,5}{0,01}\right)^2 \approx 16\,589,4$

A amostra terá de ter dimensão 16 590.

Exercícios tipo exame

Pág. 234

1.1. Para estimar a proporção populacional podemos recorrer à proporção amostral.

$\hat{p} = 12\% + 24\% = 36\%$

Estima-se que 36% dos eleitores se enquadram nos níveis 1 ou 2.

1.2. $$\left]0,08-1,645\times\sqrt{\frac{0,08(1-0,08)}{1010}};\ 0,08+1,645\times\sqrt{\frac{0,08(1-0,08)}{1010}}\right[$$

$\approx]0,066\ ;\ 0,094[$

Com uma confiança de 90%, estima-se que a proporção de eleitores que se consideram muito informados está compreendida entre $6,6\%$ e $9,4\%$, aproximadamente.

1.3. A margem de erro é igual a metade da amplitude do intervalo de confiança.

Para $\hat{p} = 0,6$, com um nível de confiança de 95%, tem-se:

***n* = 100**

$$\left]0,6-1,96\times\sqrt{\frac{0,6(1-0,6)}{100}};0,6+1,96\times\sqrt{\frac{0,6(1-0,6)}{100}}\right[\approx$$

$\approx]0,504\ ;\ 0,696[$

Margem de erro: $\frac{0,696-0,504}{2}=0,096$

***n* = 1000**

$$\left]0,6-1,96\times\sqrt{\frac{0,6(1-0,6)}{1000}};0,6+1,96\times\sqrt{\frac{0,6(1-0,6)}{1000}}\right[\approx$$

$\approx]0,570\ ;\ 0,630[$

Margem de erro: $\frac{0,630-0,570}{2}=0,03$

***n* = 5000**

$$\left]0,6-1,96\times\sqrt{\frac{0,6(1-0,6)}{5000}};0,6+1,96\times\sqrt{\frac{0,6(1-0,6)}{5000}}\right[\approx$$

$\approx]0,586\ ;\ 0,614[$

Margem de erro: $\frac{0,614-0,586}{2}=0,014$

Como se pode observar, mantendo a confiança, a margem de erro diminui à medida que a dimensão da amostra aumenta.

Pág. 235

2.1. Podemos recorrer à média amostral para estimar o valor médio.

Assim:

$$\overline{x}=\frac{32\times0+252\times1+728\times2+541\times3+494\times4+58\times5+25\times6}{2130}$$

$\approx 2,7$

Estima-se que em Portugal existe, em média, $2,7$ dispositivos com ligação à internet por habitação.

2.2. Em primeiro lugar temos de determinar o desvio-padrão amostral s.

$$s=\sqrt{\frac{233,28+728,28+356,72+48,69+834,86+306,82+272,25}{2129}}$$

$\approx 1,142\ 89$

Intervalo de confiança:

$$\left]2,7-1,96\times\frac{1,14289}{\sqrt{2130}};2,7+1,96\times\frac{1,14289}{\sqrt{2130}}\right[\approx]2,65;\ 2,75[$$

Assim, com uma confiança de 95%, estima-se que o número médio de dispositivos com ligação à Internet, por habitação, em Portugal, no ano 2015, esteja compreendido entre $2,65$ e $2,75$, aproximadamente.

Logo, com 95% de confiança, não haverá motivos para duvidar do aumento referido.

Pág. 236

3.1. $79 + 252 + 520 = 851$

851 alunos

3.2. $$p=\frac{11\ 827+42\ 334+624}{69\ 421+62\ 177+58\ 529}=\frac{54\ 785}{190\ 127}\approx 0,2881$$

Aproximadamente, $28,8\%$ dos alunos.

3.3. A: ter 18 anos $\qquad$ B: frequenta o 11.º ano

$$P(A\mid B)=\frac{P(A\cap B)}{P(B)}=\frac{\frac{5160}{190\ 127}}{\frac{62\ 177}{190\ 127}}=\frac{5160}{62\ 177}\approx 0,083$$

3.4. $$\left]0,2881-2,576\times\sqrt{\frac{0,2881(1-0,2881)}{1200}};\ 0,2881+2,576\times\sqrt{\frac{0,2881(1-0,2881)}{1200}}\right[$$

$\approx]0,25;\ 0,32[$

Com 99% de confiança, estima-se que a proporção de alunos com 16 anos matriculados em CH nas regiões autónomas está compreendida entre 25% e 32%, aproximadamente.

Pág. 237

4.1. Admitamos que X representa o vencimento bruto mensal dos funcionários.

Sabe-se que $X \sim N(1410;\ 350)$, logo o teorema do limite central garante que a distribuição de amostragem da média pode ser aproximada por uma distribuição normal com valor médio μ e desvio-padrão $\frac{\sigma}{\sqrt{n}}$.

Assim, $\overline{X} \sim N\left(1410;\ \frac{350}{\sqrt{80}}\right)$, ou seja, $\overline{X} \sim N(1410;\ 39,13)$.

Valor médio: 1410 €

Desvio-padrão: 39,13 €

4.2. $P\left(\overline{X} \leq 1500\right) = 0,5 + P\left(1410 \leq \overline{X} \leq 1500\right) \approx$

$\approx 0,5 + 0,489 = 0,989 = 98,9\%$

Probabilidade pedida: aproximadamente, $98,9\%$.

4.3. Pretende-se determinar $P(|\overline{X} - \mu| \leq 90)$.

Seja $Z \sim N(0, 1)$, então:

$$Z = \frac{\overline{X} - 1410}{39,13} \Leftrightarrow \overline{X} = 39,13Z + 1410$$

Assim:

$P(|\overline{X} - \mu| \leq 90) = P(|39,13Z + 1410 - 1410| \leq 90) =$

$= P(-90 \leq 39,13Z \leq 90) = P\left(\frac{-90}{39,13} \leq Z \leq \frac{90}{39,13}\right)$

$\approx 0,9786$

Probabilidade pedida: aproximadamente, $0,9786$.