III. Introdução à inferência estatística

Capítulo 4 Introdução à inferência estatística

Check pág. 97

1. Esta amostra é enviesada porque as 80 pessoas foram escolhidas apenas nas imediações de um ginásio. Estas pessoas têm maior probabilidade de se preocupar com saúde e alimentação, o que pode não representar os hábitos da população da freguesia em geral. Assim, os resultados obtidos podem não refletir a realidade de toda a comunidade.

2. Como o inquérito é feito presencialmente por apenas dois alunos, é provável que eles escolham colegas seus ou pessoas mais acessíveis, como alunos da mesma turma ou ano, o que pode não representar toda a população escolar e criar um enviesamento na amostra. De facto, a forma como os alunos são abordados pode limitar a variedade da amostra e pode não representar a preferência de toda a escola em relação às redes sociais.

Prova dos 9 pág. 101

Apresentamos algumas sugestões de abordagem para cada atividade.

Atividade 1

Relatório: seleção aleatória de clientes para estudo
Objetivo do estudo: selecionar 25 clientes de forma aleatória entre os 753 que fizeram compras nos últimos dois meses, para um estudo sobre a sua satisfação.
O que é uma tabela de números aleatórios? É uma lista de números gerados de forma totalmente ao acaso. Serve para fazer seleções imparciais e justas, como num sorteio, garantindo que todos têm a mesma probabilidade de ser escolhidos.

Exemplo de tabela de números aleatórios:

456	982	245	319	631	134	723	009	387	889
512	210	301	654	078	399	945	145	222	067
812	300	744	038	120	951	619	011	017	753

Critério: aceitam-se apenas números entre 001 e 753 (limite da população).

Processo de seleção:

1. Ordenar os nomes dos clientes alfabeticamente e atribuir um número a cada um, de 001 a 753, conforme a posição na lista.
2. Lê-se a tabela da esquerda para a direita, linha a linha, de cima para baixo.
3. Rejeitam-se os números superiores a 753.
4. Selecionam-se os primeiros 25 números válidos e diferentes.

Clientes selecionados: 456, 245, 319, 631, 134, 723, 009, 387, 512, 210, 301, 654, 078, 399, 145, 222, 067, 300, 744, 038, 120, 619, 011, 017, 753.

Conclusão: este processo garante uma seleção justa e aleatória, desde que todos os clientes tenham a mesma probabilidade de serem selecionados. A forma como são numerados pode variar, desde que o critério seja neutro e consistente.

Atividade 2

Os números aleatórios são essenciais na criptografia, a ciência que se ocupa em proteger informações, pois ajudam a criar chaves seguras — como palavras-passe muito difíceis de adivinhar —, que servem para codificar e descodificar mensagens.

Um dos usos mais antigos e famosos foi com a Máquina Enigma, usada pela Alemanha na II Guerra Mundial. Esta máquina criava mensagens codificadas com base em configurações que mudavam todos os dias, gerando combinações quase impossíveis de prever. A quebra do código da Enigma pelos Aliados foi um dos fatores que ajudou a antecipar o fim da guerra.

Hoje, os números aleatórios continuam a ser usados em muitas áreas da tecnologia, como nas comunicações seguras na internet, nos cartões bancários e até em aplicações como o WhatsApp.

Quanto mais imprevisíveis forem os números gerados, mais difícil será descobrir ou decifrar a informação protegida. Por isso, a aleatoriedade é um dos pilares fundamentais da segurança digital moderna.

Atividade 3

A amostragem é uma parte essencial dos estudos estatísticos, pois permite analisar uma parte da população para tirar conclusões sobre o todo. Existem diferentes tipos de amostragem, sendo duas das mais comuns a amostragem sistemática e a amostragem aleatória simples.

A amostragem sistemática consiste em escolher elementos da população de forma regular, depois de determinar aleatoriamente um ponto de partida. Por exemplo, se quisermos escolher 10 pessoas de um grupo de 100, podemos selecionar uma a cada 10, como a 5ª, a 15ª, a 25ª, e assim por diante. Este método é útil quando os elementos da população estão listados de forma aleatória ou ordenada, como em listas telefónicas ou registos de clientes.

A principal vantagem da amostragem sistemática é a sua simplicidade e rapidez na implementação. É fácil de aplicar e pode proporcionar uma boa distribuição dos elementos da população. No entanto, pode tornar-se problemática se houver padrões repetitivos nos dados, pois isso pode introduzir viés na seleção.

Por outro lado, a amostragem aleatória simples baseia-se na seleção de elementos de forma totalmente aleatória, garantindo que todos têm exatamente a mesma probabilidade de ser escolhidos. É considerada mais justa e imparcial, sendo ideal para estudos onde a representatividade é crucial. Contudo, este método pode ser mais demorado e requer ferramentas para a sua implementação, como geradores de números aleatórios ou tabelas de números aleatórios, o que pode dificultar a sua aplicação em contextos com poucos recursos.

Em conclusão, tanto a amostragem sistemática como a aleatória simples têm as suas vantagens e desvantagens. A escolha entre uma e outra depende do tipo de estudo, da organização da população e dos recursos disponíveis. A sistemática é ideal para situações em que é necessário rapidez e eficácia, enquanto a aleatória simples é preferida quando se procura total imparcialidade.

Check pág. 103

1. O valor 53 kg representa o parâmetro e o valor 49,5 kg representa a estatística.

2. O valor 1,6 horas representa a estatística.

3. O valor 83% representa a estatística.

4. O valor 72,5% representa o parâmetro.

Check pág. 104

1. $\frac{54}{90} \times 100 = 60\%$

2. $100\% - 45\% = 55\%$ não praticam desporto regularmente.
$880 \times 0{,}55 = 484$ alunos

3. O parâmetro é 45%, que representa a proporção de alunos da escola que praticam desporto regularmente, e a estatística é 60%, que representa a proporção dos alunos da amostra que praticam desporto regularmente.

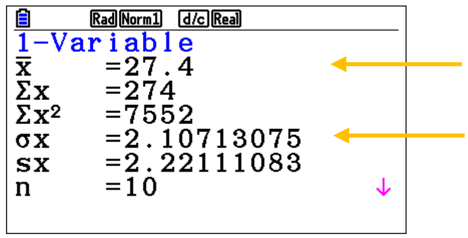
Prova dos 9 pág. 105

1.1 População: todos os portugueses adultos. Amostra: os 1102 portugueses em idade adulta.

1.2 Representa a estatística, uma vez que foi obtido a partir de uma amostra.

1.3 A veracidade da afirmação pode depender da forma como foi recolhida a amostra.

2.1 e **2.2** Os valores da média, 27,4 °C, e do desvio padrão, 2,11 °C, podem ser obtidos a partir da calculadora gráfica no modo estatístico:



Check pág. 111

1. e **2.** São 216 amostras ($6^3 = 216$). Na tabela seguinte encontram-se todas as amostras possíveis, bem como a média de cada uma:

Amostras	$\bar{x}$	Amostras	$\bar{x}$	Amostras	$\bar{x}$	Amostras	$\bar{x}$	Amostras	$\bar{x}$	Amostras	$\bar{x}$
C_1, C_1, C_1	28	C_2, C_1, C_1	27,(3)	C_3, C_1, C_1	26,(6)	C_4, C_1, C_1	27,(3)	C_5, C_1, C_1	28,(6)	C_6, C_1, C_1	26
C_1, C_1, C_2	27,(3)	C_2, C_1, C_2	26,(6)	C_3, C_1, C_2	26	C_4, C_1, C_2	26,(6)	C_5, C_1, C_2	26	C_6, C_1, C_2	25,(3)
C_1, C_1, C_3	26,(6)	C_2, C_1, C_3	26	C_3, C_1, C_3	26	C_4, C_1, C_3	26	C_5, C_1, C_3	27,(3)	C_6, C_1, C_3	24,(6)
C_1, C_1, C_4	27,(3)	C_2, C_1, C_4	26,(6)	C_3, C_1, C_4	26	C_4, C_1, C_4	26,(6)	C_5, C_1, C_4	28	C_6, C_1, C_4	25,(3)
C_1, C_1, C_5	28,(6)	C_2, C_1, C_5	26	C_3, C_1, C_5	27,(3)	C_4, C_1, C_5	28	C_5, C_1, C_5	29,3	C_6, C_1, C_5	26,(6)
C_1, C_1, C_6	26	C_2, C_1, C_6	25,(3)	C_3, C_1, C_6	24,(6)	C_4, C_1, C_6	25,(3)	C_5, C_1, C_6	26,(6)	C_6, C_1, C_6	24
C_1, C_2, C_1	27,(3)	C_2, C_2, C_1	26,(6)	C_3, C_2, C_1	26	C_4, C_2, C_1	26,(6)	C_5, C_2, C_1	26	C_6, C_2, C_1	25,(3)
C_1, C_2, C_2	26,(6)	C_2, C_2, C_2	26	C_3, C_2, C_2	25,(3)	C_4, C_2, C_2	26	C_5, C_2, C_2	27,(3)	C_6, C_2, C_2	24,(6)
C_1, C_2, C_3	26	C_2, C_2, C_3	25,(3)	C_3, C_2, C_3	24,(6)	C_4, C_2, C_3	25,(3)	C_5, C_2, C_3	26,(6)	C_6, C_2, C_3	24
C_1, C_2, C_4	26,(6)	C_2, C_2, C_4	26	C_3, C_2, C_4	25,(3)	C_4, C_2, C_4	26	C_5, C_2, C_4	27,(3)	C_6, C_2, C_4	24,(6)
C_1, C_2, C_5	28	C_2, C_2, C_5	27,(3)	C_3, C_2, C_5	26,(6)	C_4, C_2, C_5	27,(3)	C_5, C_2, C_5	28,(6)	C_6, C_2, C_5	26
C_1, C_2, C_6	25,(3)	C_2, C_2, C_6	24,(6)	C_3, C_2, C_6	24	C_4, C_2, C_6	24,(6)	C_5, C_2, C_6	26	C_6, C_2, C_6	23,(3)
C_1, C_3, C_1	26,(6)	C_2, C_3, C_1	26	C_3, C_3, C_1	26	C_4, C_3, C_1	26	C_5, C_3, C_1	27,(3)	C_6, C_3, C_1	24,(6)
C_1, C_3, C_2	26	C_2, C_3, C_2	25,(3)	C_3, C_3, C_2	24,(6)	C_4, C_3, C_2	25,(3)	C_5, C_3, C_2	26,(6)	C_6, C_3, C_2	24
C_1, C_3, C_3	26	C_2, C_3, C_3	24,(6)	C_3, C_3, C_3	24	C_4, C_3, C_3	24,(6)	C_5, C_3, C_3	26	C_6, C_3, C_3	23,(3)

C_1, C_3, C_4	26	C_2, C_3, C_4	25,(3)	C_3, C_3, C_4	24,(6)	C_4, C_3, C_4	25,(3)	C_5, C_3, C_4	26,(6)	C_6, C_3, C_4	24
C_1, C_3, C_5	27,(3)	C_2, C_3, C_5	26,(6)	C_3, C_3, C_5	26	C_4, C_3, C_5	26,(6)	C_5, C_3, C_5	28	C_6, C_3, C_5	25,(3)
C_1, C_3, C_6	24,(6)	C_2, C_3, C_6	24	C_3, C_3, C_6	23,(3)	C_4, C_3, C_6	24	C_5, C_3, C_6	25,(3)	C_6, C_3, C_6	22,(6)
C_1, C_4, C_1	27,(3)	C_2, C_4, C_1	26,(6)	C_3, C_4, C_1	26	C_4, C_4, C_1	26,(6)	C_5, C_4, C_1	28	C_6, C_4, C_1	25,(3)
C_1, C_4, C_2	26,(6)	C_2, C_4, C_2	26	C_3, C_4, C_2	25,(3)	C_4, C_4, C_2	26	C_5, C_4, C_2	27,(3)	C_6, C_4, C_2	24,(6)
C_1, C_4, C_3	26	C_2, C_4, C_3	25,(3)	C_3, C_4, C_3	24,(6)	C_4, C_4, C_3	25,(3)	C_5, C_4, C_3	26,(6)	C_6, C_4, C_3	24
C_1, C_4, C_4	26,(6)	C_2, C_4, C_4	26	C_3, C_4, C_4	25,(3)	C_4, C_4, C_4	26	C_5, C_4, C_4	27,(3)	C_6, C_4, C_4	24,(6)
C_1, C_4, C_5	28	C_2, C_4, C_5	27,(3)	C_3, C_4, C_5	26,(6)	C_4, C_4, C_5	27,(3)	C_5, C_4, C_5	27,(3)	C_6, C_4, C_5	26
C_1, C_4, C_6	25,(3)	C_2, C_4, C_6	24,(6)	C_3, C_4, C_6	24	C_4, C_4, C_6	24,(6)	C_5, C_4, C_6	26	C_6, C_4, C_6	23,(3)
C_1, C_5, C_1	28,(6)	C_2, C_5, C_1	26	C_3, C_5, C_1	27,(3)	C_4, C_5, C_1	28	C_5, C_5, C_1	29,3	C_6, C_5, C_1	26,(6)
C_1, C_5, C_2	26	C_2, C_5, C_2	25,(3)	C_3, C_5, C_2	26,(6)	C_4, C_5, C_2	27,(3)	C_5, C_5, C_2	28,(6)	C_6, C_5, C_2	26
C_1, C_5, C_3	27,(3)	C_2, C_5, C_3	26,(6)	C_3, C_5, C_3	26	C_4, C_5, C_3	26,(6)	C_5, C_5, C_3	28	C_6, C_5, C_3	25,(3)
C_1, C_5, C_4	28	C_2, C_5, C_4	27,(3)	C_3, C_5, C_4	26,(6)	C_4, C_5, C_4	27,(3)	C_5, C_5, C_4	27,(3)	C_6, C_5, C_4	26
C_1, C_5, C_5	29,3	C_2, C_5, C_5	28,(6)	C_3, C_5, C_5	28	C_4, C_5, C_5	27,(3)	C_5, C_5, C_5	30	C_6, C_5, C_5	27,(3)
C_1, C_5, C_6	26,(6)	C_2, C_5, C_6	26	C_3, C_5, C_6	25,(3)	C_4, C_5, C_6	26	C_5, C_5, C_6	27,(3)	C_6, C_5, C_6	24,(6)
C_1, C_6, C_1	26	C_2, C_6, C_1	25,(3)	C_3, C_6, C_1	24,(6)	C_4, C_6, C_1	25,(3)	C_5, C_6, C_1	26,(6)	C_6, C_6, C_1	26
C_1, C_6, C_2	25,(3)	C_2, C_6, C_2	24,(6)	C_3, C_6, C_2	24	C_4, C_6, C_2	24,(6)	C_5, C_6, C_2	26	C_6, C_6, C_2	23,(3)
C_1, C_6, C_3	24,(6)	C_2, C_6, C_3	24	C_3, C_6, C_3	23,(3)	C_4, C_6, C_3	24	C_5, C_6, C_3	25,(3)	C_6, C_6, C_3	22,(6)
C_1, C_6, C_4	25,(3)	C_2, C_6, C_4	24,(6)	C_3, C_6, C_4	24	C_4, C_6, C_4	24,(6)	C_5, C_6, C_4	26	C_6, C_6, C_4	23,(3)
C_1, C_6, C_5	26,(6)	C_2, C_6, C_5	26	C_3, C_6, C_5	25,(3)	C_4, C_6, C_5	26	C_5, C_6, C_5	27,(3)	C_6, C_6, C_5	24,(6)
C_1, C_6, C_6	24	C_2, C_6, C_6	23,(3)	C_3, C_6, C_6	22,(6)	C_4, C_6, C_6	23,(3)	C_5, C_6, C_6	24,(6)	C_6, C_6, C_6	22

3. Para os 216 casos possíveis, façamos a contagem dos casos favoráveis para cada um dos valores de $\bar{x}$:

$\boldsymbol{x_i}$	22	22,(6)	23,(3)	24	24,(6)	25,(3)	26	26,(6)	27,(3)	28	28,(6)	29,(3)	30
$\boldsymbol{P(\bar{X} = x_i)}$	$\frac{1}{216}$	$\frac{3}{216}$	$\frac{9}{216}$	$\frac{16}{216}$	$\frac{27}{216}$	$\frac{33}{216}$	$\frac{38}{216}$	$\frac{33}{216}$	$\frac{27}{216}$	$\frac{16}{216}$	$\frac{9}{216}$	$\frac{3}{216}$	$\frac{1}{216}$

Com o auxílio, por exemplo, da calculadora gráfica, chegamos aos valores:

$$E(\bar{X}) = 26 \ \text{ e } \ \sigma_{\bar{X}} \approx 1{,}491$$

4. O estimador é não enviesado, uma vez que a média da distribuição de amostragem coincide com o parâmetro a estimar.

5. O desvio padrão da distribuição de amostragem da média é aproximadamente 1,491, valor menor do que no caso de amostras de dimensão 2 (que era de aproximadamente 1,826). Podemos dizer que, ao aumentar a dimensão da amostra, diminuímos a variabilidade da estatística.

1. $4^2 = 16$ amostras

2. Cálculo das médias das amostras:

Amostras	$\bar{x}$	Amostras	$\bar{x}$	Amostras	$\bar{x}$	Amostras	$\bar{x}$
(2, 2)	2	(4, 2)	3	(7, 2)	4,5	(9, 2)	5,5
(2, 4)	3	(4, 4)	4	(7, 4)	5,5	(9, 4)	6,5
(2, 7)	4,5	(4, 7)	5,5	(7, 7)	7	(9,7)	8
(2, 9)	5,5	(4, 9)	6,5	(7, 9)	8	(9, 9)	9

e a distribuição de amostragem:

$\boldsymbol{x_i}$	2	3	4	4,5	5,5	6,5	7	8	9
$\boldsymbol{P(\bar{X} = x_i)}$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{4}{16}$	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{1}{16}$

3. $E(\bar{X}) = \frac{1}{16} \times 2 + \frac{2}{16} \times 3 + \frac{1}{16} \times 4 + \frac{2}{16} \times 4{,}5 \times \frac{4}{16} \times 5{,}5 + \frac{2}{16} \times 6{,}5 + \frac{1}{16} \times 7 + \frac{2}{16} \times 8 + \frac{1}{16} \times 9 = 5{,}5$

$\mu = \frac{2+4+7+9}{4} = 5{,}5$

Como os valores são iguais, o estimador não é enviesado.

4. $\sigma_{\bar{X}} \approx 1{,}904$ e $\sigma \approx 2{,}693$ (valores obtidos na calculadora gráfica no modo estatístico)

5. $\frac{\sigma}{\sqrt{2}} \approx \frac{2{,}693}{\sqrt{2}} \approx 1{,}904 \approx \sigma_{\bar{X}}$

1.1 $4^3 = 64$ amostras

1.2 Cálculo das médias das amostras:

Amostras	$\overline{x}$	Amostras	$\overline{x}$	Amostras	$\overline{x}$	Amostras	$\overline{x}$	Amostras	$\overline{x}$	Amostras	$\overline{x}$	Amostras	$\overline{x}$	Amostras	$\overline{x}$
(2, 2, 2)	2	(2, 7, 2)	$\frac{11}{3}$	(4, 2, 2)	$\frac{8}{3}$	(4, 7, 2)	$\frac{13}{3}$	(7, 2, 2)	$\frac{11}{3}$	(7, 7, 2)	$\frac{16}{3}$	(9, 2, 2)	$\frac{13}{3}$	(9, 7, 2)	6
(2, 2, 4)	$\frac{8}{3}$	(2, 7, 4)	$\frac{13}{3}$	(4, 2, 4)	$\frac{10}{3}$	(4, 7, 4)	5	(7, 2, 4)	$\frac{13}{3}$	(7, 7, 4)	6	(9, 2, 4)	5	(9, 7, 4)	$\frac{20}{3}$
(2, 2, 7)	$\frac{11}{3}$	(2, 7, 7)	$\frac{16}{3}$	(4, 2, 7)	$\frac{13}{3}$	(4, 7, 7)	6	(7, 2, 7)	$\frac{16}{3}$	(7, 7, 7)	7	(9, 2, 7)	6	(9, 7, 7)	$\frac{23}{3}$
(2, 2, 9)	$\frac{13}{3}$	(2, 7, 9)	6	(4, 2, 9)	5	(4, 7, 9)	$\frac{20}{3}$	(7, 2, 9)	6	(7, 7, 9)	$\frac{23}{3}$	(9, 2, 9)	$\frac{13}{3}$	(9, 7, 9)	$\frac{25}{3}$
(2, 4, 2)	$\frac{8}{3}$	(2, 9, 2)	$\frac{13}{3}$	(4, 4, 2)	$\frac{10}{3}$	(4, 9, 2)	5	(7, 4, 2)	$\frac{13}{3}$	(7, 9, 2)	6	(9, 4, 2)	5	(9, 9, 2)	$\frac{20}{3}$
(2, 4, 4)	$\frac{10}{3}$	(2, 9, 4)	5	(4, 4, 4)	4	(4, 9, 4)	$\frac{17}{3}$	(7, 4, 4)	5	(7, 9, 4)	$\frac{20}{3}$	(9, 4, 4)	$\frac{17}{3}$	(9, 9, 4)	$\frac{22}{3}$
(2, 4, 7)	$\frac{13}{3}$	(2, 9, 7)	6	(4, 4, 7)	5	(4, 9, 7)	$\frac{20}{3}$	(7, 4, 7)	6	(7, 9, 7)	$\frac{23}{3}$	(9, 4, 7)	$\frac{20}{3}$	(9, 9, 7)	$\frac{25}{3}$
(2, 4, 9)	5	(2, 9, 9)	$\frac{20}{3}$	(4, 4, 9)	$\frac{17}{3}$	(4, 9, 9)	$\frac{22}{3}$	(7, 4, 9)	$\frac{20}{3}$	(7, 9, 9)	$\frac{25}{3}$	(9, 4, 9)	$\frac{22}{3}$	(9, 9, 9)	9

e a distribuição de amostragem:

$\boldsymbol{x_i}$	2	$\frac{8}{3}$	$\frac{10}{3}$	$\frac{11}{3}$	4	$\frac{13}{3}$	5	$\frac{16}{3}$	$\frac{17}{3}$	6	$\frac{20}{3}$	7	$\frac{22}{3}$	$\frac{23}{3}$	$\frac{25}{3}$	9
$\boldsymbol{P(\overline{X} = x_i)}$	$\frac{1}{64}$	$\frac{3}{64}$	$\frac{3}{64}$	$\frac{3}{64}$	$\frac{1}{64}$	$\frac{9}{64}$	$\frac{9}{64}$	$\frac{3}{64}$	$\frac{3}{64}$	$\frac{9}{64}$	$\frac{9}{64}$	$\frac{1}{64}$	$\frac{3}{64}$	$\frac{3}{64}$	$\frac{3}{64}$	$\frac{1}{64}$

1.3

$$E(\overline{X}) = \frac{1}{64} \times (2+4+7+9) + \frac{3}{64} \times \left(\frac{8}{3} + \frac{10}{3} + \frac{11}{3} + \frac{16}{3} + \frac{17}{3} + \frac{22}{3} + \frac{23}{3} + \frac{25}{3}\right) + \frac{9}{64} \times \left(\frac{13}{3} + 5 + 6 + \frac{20}{3}\right) = 5{,}5$$

$\mu = \frac{2+4+7+9}{4} = 5{,}5$

Como os valores são iguais, o estimador não é enviesado.

1.4 $\sigma_{\overline{X}} \approx 1{,}555$ e $\sigma \approx 2{,}693$ (valores obtidos a partir da calculadora gráfica no modo estatístico)

1.5 $\frac{\sigma}{\sqrt{3}} \approx \frac{2{,}693}{\sqrt{3}} \approx 1{,}555 \approx \sigma_{\overline{X}}$

Check pág. 119

1. $\mu_{\bar{X}} = \mu = 10$ horas

$\sigma_{\bar{X}} = \frac{4}{\sqrt{40}} \approx 0{,}632$ horas

2. A dimensão da amostra é 40 (maior do que 30), logo, pelo teorema do limite central, podemos afirmar que a distribuição de amostragem da média pode ser aproximada por uma distribuição normal com um valor médio 10 e desvio padrão 0,632.

Seja $U \sim N(0{,}1)$, então, $U = \frac{\bar{X}-10}{0{,}632} \Leftrightarrow \bar{X} = 0{,}632U + 10$.

Vem que:

$P(8 < \bar{X} < 11) = P(-3{,}16 < U < 1{,}58) = \phi(1{,}58) - [1 - \phi(3{,}16)] = 0{,}9429 - 1 + 0{,}9992 = 0{,}9421$

(Valores da tabela da distribuição normal *standard*, página 62 do Manual.)

Assim, a probabilidade de, na amostra, o número de horas de estudo estar entre 8 e 11 horas é 94,21%.

3. $P(|\bar{X} - \mu| \leq 0{,}5) = P(|U| \leq 0{,}79) = 2 \times \phi(0{,}79) - 1 = 2 \times 0{,}7852 - 1 = 0{,}5704$

A probabilidade pedida é 57,04%.

Prova dos 9 pág. 119

1.1 $E(\bar{X}) = \mu = 950$ euros

$\sigma_{\bar{X}} = \frac{250}{\sqrt{84}} \approx 27{,}277$ euros

1.2.1 Seja $U \sim N(0{,}1)$, então, $U = \frac{\bar{X}-950}{27{,}277} \Leftrightarrow \bar{X} = 27{,}277U + 950$.

Vem que:

$$P(\bar{X} \leq 1000) = P(U \leq 1{,}83) = \phi(1{,}83) = 0{,}9664$$

Assim, a probabilidade de, na amostra, o gasto com as férias de verão ser inferior ou igual a 1000 euros é 96,64%.

1.2.2 $P(860 < \bar{X} < 1020) = P(-3{,}30 < U < 2{,}57) = \phi(2{,}57) - [1 - \phi(3{,}30)] =$

$= 0{,}9949 - 1 + 0{,}9995 = 0{,}9944$

A probabilidade pedida é 99,44%.

2.1 O valor médio da distribuição de amostragem é $E(\bar{X}) = 15{,}3$. O desvio padrão amostral é:

$$\sigma_{\bar{X}} = \frac{7{,}02}{\sqrt{60}} \approx 0{,}906$$

2.2 Como a dimensão da amostra é maior do que 30 (é igual a 60), então, pelo teorema do limite central, podemos afirmar que a distribuição de amostragem da média pode ser aproximada a uma distribuição normal com valor médio de 15,3 e desvio padrão de 0,906.

Seja $U \sim N(0, 1)$.

Então:

$$U=\frac{\bar{X}-15{,}3}{0{,}906}\Leftrightarrow \bar{X}=0{,}906U+15{,}3$$

E, assim:

$$\begin{aligned}P(13<\bar{X}<15)&=P(13<0{,}906U+15{,}3<15)=P(-2{,}539<U<-0{,}331)=\\&=1-\phi(0{,}33)-1+\phi(2{,}54)=-0{,}6293+0{,}9945=\\&=0{,}3652\end{aligned}$$

A probabilidade de o número médio de defeitos (na amostra) estar entre 13 e 15 defeitos por cada 100 metros de tecido é de aproximadamente 36,52%.

2.3 $P(|\bar{X}-\mu|\leq 0{,}5)=P(|0{,}906U+15{,}3-15{,}3|\leq 0{,}5)=P(|U|<0{,}552)=$

$$\begin{aligned}&=P(-0{,}552<U<0{,}552)=2\times\phi(0{,}55)-1=\\&=2\times 0{,}7088-1=\\&=0{,}4176\end{aligned}$$

A probabilidade pedida é de aproximadamente 41,76%.

Check pág. 122

1. $\mu=120\text{ mm}$

Um intervalo de confiança de 90% para o parâmetro μ é de forma:

$$\left]\bar{x}-1{,}645\times\frac{\sigma}{\sqrt{n}};\ \bar{x}+1{,}645\times\frac{\sigma}{\sqrt{n}}\right[=\left]120-1{,}645\times\frac{10}{\sqrt{50}};\ 120+1{,}645\times\frac{10}{\sqrt{50}}\right[=]117{,}67;\ 122{,}33[$$

2. Para um nível de confiança de 90%, $z\approx 1{,}645$.

Então, um intervalo de confiança de 90% para o parâmetro μ é da forma:

$$\left]\bar{x}-1{,}645\times\frac{\sigma}{\sqrt{n}};\bar{x}+1{,}645\times\frac{\sigma}{\sqrt{n}}\right[=\left]18-1{,}645\times\frac{4}{\sqrt{40}};18+1{,}645\times\frac{4}{\sqrt{40}}\right[=]16{,}960;19{,}040[$$

Check pág. 123

$n=1000$ $\quad\bar{x}=75$ $\quad s=10$

Um intervalo de confiança para μ com 95% de confiança é dado por:

$$\left]75-1{,}96\times\frac{10}{\sqrt{1000}};\ 75+1{,}96\times\frac{10}{\sqrt{1000}}\right[=]74{,}38;\ 75{,}62[$$

Check pág. 124

$n = 200 \qquad \bar{x} = 110 \qquad s = 15$

1. Um intervalo de confiança para μ com 99% de confiança é dado por:

$$\left]110 - 2{,}576 \times \frac{15}{\sqrt{200}};\ 110 + 2{,}576 \times \frac{15}{\sqrt{200}}\right[=]107{,}27;\ 112{,}73[$$

2. Não, porque a amostra tem dimensão superior a 30.

Prova dos 9 pág. 125

1.1 $n = 50$, $\bar{x} = 3{,}5\ \text{h}$ e $\sigma = 1{,}2\ \text{h}$

Para um nível de confiança de 90%, $z \approx 1{,}645$.

Então, um intervalo de confiança de 90% para o parâmetro μ é da forma:

$$\left]\bar{x} - 1{,}645 \times \frac{\sigma}{\sqrt{n}}; \bar{x} + 1{,}645 \times \frac{\sigma}{\sqrt{n}}\right[= \left]3{,}5 - 1{,}645 \times \frac{1{,}2}{\sqrt{50}}; 3{,}5 + 1{,}645 \times \frac{1{,}2}{\sqrt{50}}\right[=]3{,}22; 3{,}78[$$

1.2 Se recolhermos dados em 100 amostras, apenas 10 delas conduzirão a intervalos que não contêm o tempo médio da população, ou existe 90% de confiança de que o verdadeiro tempo médio diário que os estudantes do Ensino Secundário passam nas redes sociais esteja entre 3,220 e 3,780 h.

2.1 $n = 36$, $\bar{x} = 45$ e $\sigma = 5$

Para um nível de confiança de 95%, $z \approx 1{,}960$.

Então, um intervalo de confiança de 95% para o parâmetro μ é da forma:

$$\left]\bar{x} - 1{,}960 \times \frac{\sigma}{\sqrt{n}}; \bar{x} + 1{,}960 \times \frac{\sigma}{\sqrt{n}}\right[= \left]45 - 1{,}960 \times \frac{5}{\sqrt{36}}; 45 + 1{,}960 \times \frac{5}{\sqrt{36}}\right[=]43{,}37; 46{,}63[$$

2.2 Se recolhermos dados em 100 amostras, apenas 5 delas conduzirão a intervalos que não contêm o tempo médio da população, ou existe 95% de confiança de que o verdadeiro tempo médio que os funcionários levam a completar a tarefa esteja entre 43,37 e 46,63 minutos.

3. $n = 400$, $\bar{x} = 22{,}8$ e $\sigma = 1{,}5$ anos

Para um nível de confiança de 99%, $z \approx 2{,}576$.

Então, um intervalo de confiança de 99% para o parâmetro μ é da forma:

$$\left]\bar{x} - 2{,}576 \times \frac{\sigma}{\sqrt{n}}; \bar{x} + 2{,}576 \times \frac{\sigma}{\sqrt{n}}\right[= \left]22{,}8 - 2{,}576 \times \frac{1{,}5}{\sqrt{400}}; 22{,}8 + 2{,}576 \times \frac{1{,}5}{\sqrt{400}}\right[=]22{,}61; 22{,}99[$$

Prova dos 9 pág. 129

1. $\mu = 75\%$, $n = 1200$, $\hat{p} = 0{,}75$ e $\sigma_{\hat{p}} = \sqrt{\frac{0{,}75(1-0{,}75)}{1200}} = 0{,}0125$

Prova dos 9 pág. 132

1. $\boldsymbol{\hat{p} = \frac{48}{150} = 0{,}32}$

O intervalo de confiança será:

$$\left]0{,}32 - 1{,}645 \times \sqrt{\frac{0{,}32(1-0{,}32)}{150}};\ 0{,}32 + 1{,}645 \times \sqrt{\frac{0{,}32(1-0{,}32)}{150}}\right[=]0{,}26;\ 0{,}38[$$

A proporção dos votos encontra-se entre os 26% e os 38%.

2. Tem-se que $n = 500$, $z = 1{,}960$ e $\hat{p} = \frac{8}{500} = 0{,}016$.

O intervalo de confiança para a proporção de livros com defeito será:

$$\left]0{,}016 - 1{,}960 \times \sqrt{\frac{0{,}016(1-0{,}016)}{500}};\ 0{,}016 + 1{,}960 \times \sqrt{\frac{0{,}016(1-0{,}016)}{500}}\right[=]0{,}005; 0{,}027[$$

3. $n = 300$ e $\hat{p} = \frac{180}{300} = 0{,}6$.

O intervalo de confiança para a proporção com 99% de confiança é da forma:

$$\left]0{,}6 - 2{,}576 \times \sqrt{\frac{0{,}6(1-0{,}6)}{300}};\ 0{,}6 + 2{,}576 \times \sqrt{\frac{0{,}6(1-0{,}6)}{300}}\right[=]0{,}53; 0{,}67[$$

Check pág. 133

1. Com uma confiança de 90%, podemos afirmar que entre 83% e 87% dos entrevistados tinham automóvel.

2. A percentagem de portugueses que experimentaram IA está entre 64,9% e 71,1%

Check pág. 135

$\varepsilon = 5 \text{ min}$ $\quad n = 50$ $\quad \sigma = 20$

$n = \left(\frac{2{,}576 \times 20}{5}\right)^2 \approx 107$

Aproximadamente 107 alunos.

Check pág. 136

$\varepsilon = 2\%$

$n = \left(\frac{1,96}{0,02}\right)^2 \times 0,35(1-0,35) \approx 2185$

Aproximadamente 2185 eleitores.

Check pág. 138

$\varepsilon = 4\%$

Tendo em conta o resultado anterior, tem-se que:

$n = \left(\frac{2,576}{0,04}\right)^2 \times 0,25 \approx 1037$

Aproximadamente 1037 pessoas.

Prova dos 9 pág. 138

1. Se a amplitude total é 0,5, então $\varepsilon = 0,25$.

$\sigma = 1$

Para uma confiança de 95%, $z = 1,960$, logo $n = \left(\frac{1,960 \times 1}{0,25}\right)^2 \approx 61,47$.

O tamanho mínimo da amostra necessário para atingir essa precisão deve ser 62 pacotes de açúcar.

2. $\varepsilon = 0,02$

Para uma confiança de 99%, $z = 2,576$.

$p = 0,5$

$n = \left(\frac{z}{\varepsilon}\right)^2 \times p \times (1-p)$, logo $n = \left(\frac{2,576}{0,02}\right)^2 \times 0,5 \times 0,5 = 4147,36$

O tamanho mínimo da amostra necessário para atingir essa precisão desejada deve ser 4148 pessoas.

3. Se a amplitude total é 2, então $\varepsilon = 1$.

$\sigma = 4$

Para uma confiança de 90%, $z = 1,645$.

$n = \left(\frac{z \times \sigma}{\varepsilon}\right)^2$, logo $n = \left(\frac{1,645 \times 4}{1}\right)^2 = 43,28$

O tamanho mínimo da amostra necessário para atingir essa precisão desejada deve ser 44 dias.

4. A margem de erro é metade da amplitude do intervalo, ou seja, $\varepsilon = \frac{270-250}{2} = 10$.

Opção (B)