4. Introdução à inferência estatística

Recorda – páginas 6 a 11

1.

a) Opção (C)

A população é o conjunto de todos os elementos sobre os quais se pretende obter informação.

Neste caso, a população é o conjunto de todas as respostas dadas pelo assistente virtual.

b) Opção (C)

A amostra é um subconjunto da população, selecionado para análise, uma vez que não é possível estudar todos os elementos.

Neste caso, a amostra é o conjunto de respostas selecionadas aleatoriamente para análise.

2. Os tempos de resposta, em segundos, registados foram:

3 5 4 6 5 7 4 6

correspondendo a uma amostra de dimensão $n = 8$.

a) A média amostral é dada por:

$$\bar{x} = \frac{3+5+4+6+5+7+4+6}{8} = \frac{40}{8} = 5$$

Logo, a média amostral do tempo de resposta é 5 segundos.

b) No caso em estudo, com $\bar{x} = 5$ e $n = 8$:

$$s = \sqrt{\frac{(3-5)^2+(5-5)^2+(4-5)^2+(6-5)^2+(5-5)^2+(7-5)^2+(4-5)^2+(6-5)^2}{7}} =$$

$$= \sqrt{\frac{4+0+1+1+0+4+1+1}{7}} = \sqrt{\frac{12}{7}} \approx 1{,}31$$

3. Podem definir-se, por exemplo, as seguintes variáveis aleatórias.

- **Soma dos valores obtidos**

  Seja X a variável aleatória que representa a soma dos números obtidos nos três lançamentos:

  $X = a + b + c$

- **Número de vezes que surge a face 6**

  Seja Y a variável aleatória que representa o número de vezes que aparece a face 6 nos três lançamentos: $Y \in \{0, 1, 2, 3\}$

- **Valor máximo obtido**

  Seja Z a variável aleatória que representa o maior valor observado nos três lançamentos:

  $Z = \text{máx.}\{a, b, c\}$

4. A variável aleatória W pode asumir os valores 0, 1, 2 e 3.

$$P(W = 0) = \frac{3 \times 3 \times 3}{6 \times 6 \times 6} = \frac{1}{8}$$

$$P(W = 1) = \frac{3 \times 3 \times 3 \times 3}{6 \times 6 \times 6} = \frac{3}{8}$$

$$P(W = 2) = \frac{3 \times 3 \times 3 \times 3}{6 \times 6 \times 6} = \frac{3}{8}$$

$$P(W = 3) = \frac{3 \times 3 \times 3}{6 \times 6 \times 6} = \frac{1}{8}$$

w	0	1	2	3
$P(W = w)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

5.

$$\mu = 0 + \frac{3}{8} + \frac{6}{8} + \frac{3}{8} = \frac{12}{8} = \frac{3}{2}$$

$$\sigma = \sqrt{(0-1{,}5)^2 \times \frac{1}{8} + (1-1{,}5)^2 \times \frac{3}{8} + (2-1{,}5)^2 \times \frac{3}{8} + (3-1{,}5)^2 \times \frac{1}{8}} =$$

$$= \sqrt{\frac{6}{8}} =$$

$$= \sqrt{\frac{3}{4}} =$$

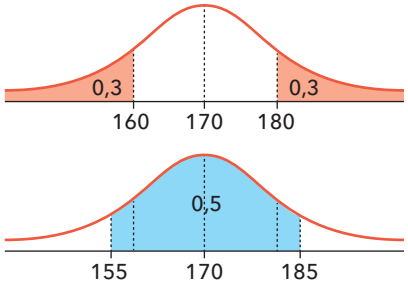
$$= \frac{\sqrt{3}}{2}$$

6. X: "Altura, em cm, de um visitante do parque temático."

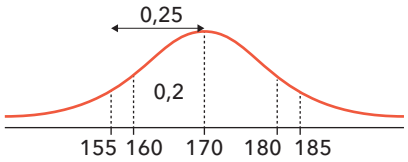
$X \sim N(170, \sigma)$.

Sabe-se que $P(X < 160) = 0{,}3$ e $P(155 < X < 185) = 0{,}5$.

Pretende-se determinar $P(155 < X < 160)$ e multiplicar este valor por 800.



Como $P(160 < X < 180) = 1 - 2 \times 0{,}3 = 0{,}4$ e $P(155 < X < 185) = 0{,}5$, então $P(160 < X < 170) = 0{,}2$ e $P(155 < X < 170) = 0{,}25$.



$P(155 < X < 160) = 0{,}25 - 0{,}2 = 0{,}05$

$0{,}05 \times 800 = 40$

O número esperado de visitantes, entre os 800 analisados, com uma altura compreendida entre 155 cm e 160 cm é 40.

7. X: "Tempo diário, em minutos, que um utilizador permanece ativo na plataforma."

$X \sim N(40, 10)$

Pretende-se determinar $P(50 < X < 60)$.

Sabe-se que, numa distribuição normal:

- cerca de 68,27% dos valores estão no intervalo $]\mu - \sigma, \mu + \sigma[=]30, 50[$;
- cerca de 95,45% dos valores estão no intervalo $]\mu - 2\sigma, \mu + 2\sigma[=]20, 60[$.

Logo, a probabilidade de um valor se situar entre os 50 e os 60 minutos é dada pela diferença entre estas duas probabilidades, repartida igualmente pelos dois lados da distribuição:

$$P(50 < X < 60) = \frac{0{,}9545 - 0{,}6827}{2} = 0{,}1359$$

A probabilidade pedida é 0,136.

Capítulo 1 – Introdução à inferência estatística

Tarefa – A sondagem – páginas 12 e 13

a) Opção (C)
b) Opção (B)
c) Opção (C)
d) Opção (B)
e) Opção (A)
f) Opção (D)
g) Opção (C)
h) Opção (B)
i) Opção (C)
j) Opção (A)

Exercícios de margem – páginas 15 a 39

1.

a) O parâmetro é a média do número de horas que todos os alunos dessa escola passam diariamente nas redes sociais.
A estatística é a média do número de horas que os alunos da amostra passam diariamente nas redes sociais, cujo valor é 2,4 horas.

b) O parâmetro é a proporção de alunos dessa universidade que vive em residências universitárias.
A estatística é a proporção de alunos da amostra que vive em residências universitárias, e é dada por $\frac{150}{500} = 0{,}3$, isto é, 30%.

2.

a) O parâmetro é a média do tempo diário de estudo de todos os alunos do 12º ano dessa escola.

b) A estatística é a média amostral.

c) $\bar{x} = \frac{2{,}5 + 3 + 1{,}5 + 4 + 2 + 3{,}5 + 2 + 4 + 3 + 2{,}5}{10} = 2{,}8$

3.

a) Número de amostras possíveis: $4 \times 4 = 16$

b)

Amostra	Média
(4, 4)	4
(4, 6)	5
(4, 8)	6
(4, 10)	7
(6, 4)	5
(6, 6)	6
(6, 8)	7
(6, 10)	8
(8, 4)	6
(8, 6)	7
(8, 8)	8
(8, 10)	9
(10, 4)	7
(10, 6)	8
(10, 8)	9
(10, 10)	10

c) Contando quantas vezes aparece cada média, obtemos:

Média	Frequência
4	1
5	2
6	3
7	4
8	3
9	2
10	1
Total	**16**

Como cada amostra tem probabilidade $\frac{1}{16}$, vem que a distribuição de amostragem da média é:

$\bar{X}$	4	5	6	7	8	9	10
Probabilidades	$\frac{1}{16}$	$\frac{2}{16} = \frac{1}{8}$	$\frac{3}{16}$	$\frac{4}{16} = \frac{1}{4}$	$\frac{3}{16}$	$\frac{2}{16} = \frac{1}{8}$	$\frac{1}{16}$

d) $\mu_{\bar{X}} = 4 \times \frac{1}{16} + 5 \times \frac{1}{8} + 6 \times \frac{3}{16} + 7 \times \frac{1}{4} + 8 \times \frac{3}{16} + 9 \times \frac{1}{8} + 10 \times \frac{1}{16} = 7$

Média da população: $\frac{4 + 6 + 8 + 10}{4} = \frac{28}{4} = 7$

Os valores são iguais.

4. Uma vez que a dimensão da amostra é superior a 30, o teorema do limite central permite afirmar que a distribuição de amostragem da média pode ser aproximada por um modelo normal, com:

- Valor médio: 22
- Desvio-padrão: $\frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{60}} \approx 0{,}39$

5.

a) Uma vez que a dimensão da amostra é superior a 30, o teorema do limite central permite afirmar que a distribuição de amostragem da média pode ser aproximada por um modelo normal, com:
Valor médio: 398
Desvio-padrão: $\frac{\sigma}{\sqrt{n}} = \frac{70}{\sqrt{100}} = 7$

b) $\bar{X} \sim N(398, 7)$

i. Pretende-se $P(391 < \bar{X} < 405)$.
Como $\mu_{\bar{x}} - \sigma_{\bar{x}} = 398 - 7 = 391$ e $\mu_{\bar{x}} + \sigma_{\bar{x}} = 398 + 7 = 405$, tem-se que $P(391 < \bar{X} < 405) \approx 0{,}6827$, isto é, aproximadamente, 68,27%.

ii. Pretende-se $P(384 < \bar{X} < 412)$.
Como $\mu_{\bar{x}} - 2\sigma_{\bar{x}} = 398 - 2 \times 7 = 384$ e $\mu_{\bar{x}} + 2\sigma_{\bar{x}} = 398 + 2 \times 7 = 412$, tem-se que $P(384 < \bar{X} < 412) \approx 0{,}9545$, isto é, aproximadamente, 95,45%.

6.

a) i. Sabemos que:
- $\bar{x} = 115$
- $\sigma = 40$
- $n = 100$

Utilizando a fórmula do intervalo de confiança, para o valor médio com uma confiança de 95%, $\left[\bar{x} - 1{,}96\frac{\sigma}{\sqrt{n}}; \bar{x} + 1{,}96\frac{\sigma}{\sqrt{n}}\right]$, e substituindo pelos respetivos valores, obtemos:

$$\left[115 - 1{,}96\frac{40}{\sqrt{100}}; 115 + 1{,}96\frac{40}{\sqrt{100}}\right] =$$

$= [107{,}16; 122{,}84]$

O intervalo de confiança de 95% para o valor médio μ das classificações dos alunos que realizaram este exame é, então, [107,16; 122,84].

ii. A margem de erro é:

$$\frac{122{,}84 - 107{,}16}{2} = \frac{15{,}68}{2} = 7{,}84$$

b) Construir um intervalo de confiança, com um nível de confiança de 95% para o valor médio, significa que, ao repetir o processo muitas vezes, em cada 100 intervalos de confiança, construídos a partir de 100 amostras da mesma dimensão, espera-se que cerca de 95 conterão o verdadeiro valor médio.

7. Sabemos que:

- $n = 150$
- $\bar{x} = 1{,}3$
- $s = 0{,}5$
- para um nível de confiança de 99%, $z \approx 2{,}576$.

Assim, substituindo os valores no intervalo $\left[\bar{x} - z\frac{s}{\sqrt{n}}; \bar{x} + z\frac{s}{\sqrt{n}}\right]$, obtemos:

$$\left[1{,}3 - 2{,}576\frac{0{,}5}{\sqrt{150}}; 1{,}3 + 2{,}576\frac{0{,}5}{\sqrt{150}}\right] =$$

$= [1{,}1948; 1{,}4052]$

O intervalo pedido, arredondando os extremos às centésimas, é, então, [1,19; 1,41].

Como o diretor afirmou que *em média, os utilizadores permanecem no ginásio cerca de 1,5 horas por sessão*, e 1,5 não pertence ao intervalo [1,19; 1,41], a Margarida tem razão para questionar a afirmação do diretor.

8. Um intervalo de 95% de confiança para o valor médio (usando o desvio-padrão amostral s) é $\left[\bar{x} - 1{,}96\frac{s}{\sqrt{n}}; \bar{x} + 1{,}96\frac{s}{\sqrt{n}}\right]$.

A amplitude do intervalo é:

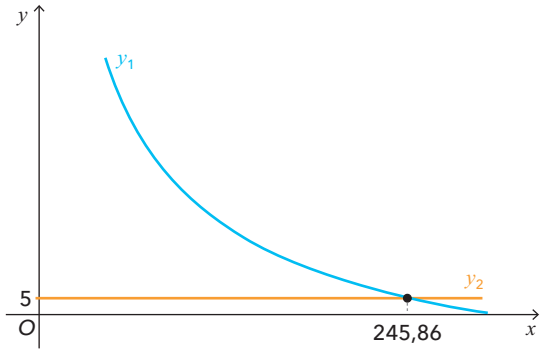
$$2 \times 1{,}96\frac{20}{\sqrt{n}} = \frac{78{,}4}{\sqrt{n}}$$

Assim, pretende-se determinar o valor de n tal que $\frac{78{,}4}{\sqrt{n}} = 5$.

Recorrendo à calculadora gráfica, e tomando x como variável independente:

$y_1 = \frac{78{,}4}{\sqrt{n}}$

$y_2 = 5$



As coordenadas do ponto de interseção são (245,86; 5).

A amostra deverá conter 246 voluntários.

9. $\hat{p} = \frac{140}{200} = 0{,}7$

10. Total de clientes: $n = 500$

Número de utilizadores que gostaram do novo sabor: 335

Assim, a proporção amostral é $\hat{p} = \frac{335}{500} = 0{,}67$.

Nível de confiança de 95%, logo $z \approx 1{,}96$.

Intervalo de confiança:

$$\left[\hat{p} - z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}; \hat{p} + z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

Substituindo os valores, obtemos o intervalo:

$$\left[0{,}67 - 1{,}96\sqrt{\frac{0{,}67(1-0{,}67)}{500}};\right.$$

$$\left.0{,}67 + 1{,}96\sqrt{\frac{0{,}67(1-0{,}67)}{500}}\right] = [0{,}6288; 0{,}7112]$$

Arredondando os limites do intervalo às centésimas, temos o intervalo de confiança pretendido: [0,63; 0,71]

11. Esta afirmação significa que o intervalo [68%, 72%] é um intervalo de confiança de 95% para a proporção pretendida, ou seja, se recolhêssemos 100 amostras diferentes, com a mesma dimensão, e construíssemos os respetivos intervalos de confiança, esperar-se-ia que, aproximadamente, 95 desses intervalos contivessem a proporção de pessoas que estão contra a construção do centro comercial e cinco intervalos não.

12.

a) $\hat{p} = 0{,}68$

Nível de confiança de 90%, logo $z \approx 1{,}645$.

Intervalo de confiança:

$$\left[\hat{p} - z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}; \hat{p} + z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

Substituindo os valores, obtemos o intervalo:

$$\left[0{,}68 - 1{,}645\sqrt{\frac{0{,}68(1-0{,}68)}{1200}};\right.$$

$$\left.0{,}68 + 1{,}645\sqrt{\frac{0{,}68(1-0{,}68)}{1200}}\right] = [0{,}6578; 0{,}7022]$$

A margem de erro é $\frac{0{,}7022 - 0{,}6578}{2} = \frac{0{,}0444}{2} =$

$= 0{,}0222$, ou seja, a margem de erro é 2,22%.

b) Substituindo $n = 1200$ por $n = 2400$, obtemos o intervalo:

$$\left[0{,}68 - 1{,}645\sqrt{\frac{0{,}68(1-0{,}68)}{2400}};\ 0{,}68 + 1{,}645\sqrt{\frac{0{,}68(1-0{,}68)}{2400}}\right] = [0{,}6643;\ 0{,}6957]$$

A margem de erro é $\frac{0{,}6957 - 0{,}6643}{2} = \frac{0{,}0314}{2} =$

$= 0{,}0157$, ou seja, a margem de erro é 1,57%.

c) A diferença entre as amplitudes dos intervalos de confiança deve-se à diferença na dimensão das amostras consideradas.

Quando a dimensão da amostra aumenta, a margem de erro diminui. Assim, ao passar de uma amostra de menor dimensão para uma amostra maior, obtém-se um intervalo de confiança mais estreito, ou seja, com menor amplitude.

Por esse motivo, o intervalo de confiança construído, com base numa amostra de maior dimensão, apresenta uma menor amplitude do que o intervalo obtido com uma amostra de menor dimensão, mantendo-se o nível de confiança.

Aprende Fazendo Global – páginas 46 a 50

1.

a) O parâmetro é a média do número de episódios de séries que todos os alunos da escola veem por semana; a estatística é a média do número de episódios de séries vistos por semana pelos alunos da amostra (6,5 episódios).

b) O parâmetro é a proporção de alunos da escola que têm subscrição paga de plataformas de *streaming*; a estatística é a proporção de alunos da amostra que têm subscrição paga, e o seu valor é $\frac{130}{200} = 0{,}65$, isto é, 65%.

2.

a) Opção (C)

$3 \times 3 = 9$

b)

Amostra	Média
(2, 2)	2
(2, 4)	3
(2, 6)	4
(4, 2)	3
(4, 4)	4
(4, 6)	5
(6, 2)	4
(6, 4)	5
(6, 6)	6

c) Como cada amostra tem probabilidade $\frac{1}{9}$, vem que a distribuição de amostragem da média é:

$\bar{X}$	2	3	4	5	6
Probabilidades	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{3}{9} = \frac{1}{3}$	$\frac{2}{9}$	$\frac{1}{9}$

d) Valor médio da distribuição de amostragem:

$$\mu_{\bar{x}} = 2 \times \frac{1}{9} + 3 \times \frac{2}{9} + 4 \times \frac{1}{3} + 5 \times \frac{2}{9} + 6 \times \frac{1}{9} = 4$$

Média da população: $\frac{2 + 4 + 6}{3} = \frac{12}{3} = 4$

Os valores são iguais.

3. Opção (B)

Uma vez que a dimensão da amostra é superior a 30, o teorema do limite central permite afirmar que a distribuição de amostragem da média pode ser aproximada por um modelo normal, com:

Valor médio: 14

Desvio-padrão: $\frac{\sigma}{\sqrt{n}} = \frac{3}{\sqrt{36}} = \frac{3}{6} = 0{,}5$

4. Opção (C)

As opções (A), (B) e (D) são verdadeiras.

Não temos informação suficiente que nos permita indicar o valor da percentagem de apoio ao partido *X*.

5. **I – a)** população

A população é o conjunto total de indivíduos sobre o qual se pretende obter informações, ou seja, todos os habitantes do concelho.

II – b) amostra

A amostra é o subconjunto de indivíduos efetivamente inquiridos para recolher dados.

III – c) estimativa pontual

A proporção de 42% obtida na amostra é uma estimativa pontual do parâmetro desconhecido *P*.

IV – b) diminuição

Mantendo o nível de confiança fixo, aumentar a dimensão da amostra reduz a amplitude do intervalo de confiança, tornando a estimativa mais precisa.

A resposta correta é: I – a); II – b); III – c); IV – b)

6.

a) Sabemos que $\bar{x} = 140$, $\sigma = 25$ e $n = 36$.

Utilizando a fórmula do intervalo de confiança, para o valor médio com uma confiança de 95%, $\left[\bar{x} - 1{,}96\frac{\sigma}{\sqrt{n}};\ \bar{x} + 1{,}96\frac{\sigma}{\sqrt{n}}\right]$, e, substituindo pelos respetivos valores, obtemos:

$$\left[140 - 1{,}96\frac{25}{\sqrt{36}};\ 140 + 1{,}96\frac{25}{\sqrt{36}}\right] =$$

$= [131{,}83;\ 148{,}17]$

O intervalo de confiança de 95% para o valor médio μ de tempo que estes jovens passam nas redes sociais é, então, [131,83; 148,17].

b) A margem de erro é $\frac{148{,}17 - 131{,}83}{2} = \frac{16{,}34}{2} = 8{,}17$.

7.

a) Opção (D)

A margem de erro do intervalo de confiança é metade da amplitude do intervalo. Assim:

$\frac{19{,}6 - 18{,}4}{2} = \frac{1{,}2}{2} = 0{,}6$

b) Sabemos que:
- $n = 64$
- $\bar{x} = 20$
- $s = 3{,}2$

Para um nível de confiança de 90%, tem-se $z = 1{,}645$.
Como a amostra tem dimensão superior a 30, podemos determinar o intervalo de confiança aplicando a fórmula $\left[\bar{x} - z\frac{s}{\sqrt{n}}\,;\,\bar{x} + z\frac{s}{\sqrt{n}}\right]$. Substituindo os valores, obtemos:

$$\left[20 - 1{,}645\frac{3{,}2}{\sqrt{64}}\,;\,20 + 1{,}645\frac{3{,}2}{\sqrt{64}}\right] =$$

$$= [19{,}342;\ 20{,}658]$$

Apresentando os extremos do intervalo de confiança arredondados às décimas, tem-se que o intervalo de confiança a 90% é [19,3; 20,7].

c) Um intervalo de 95% de confiança para o valor médio (usando o desvio-padrão amostral s) é:

$$\left[\bar{x} - 1{,}96\frac{s}{\sqrt{n}}\,;\,\bar{x} + 1{,}96\frac{s}{\sqrt{n}}\right]$$

A margem de erro é metade da amplitude do intervalo, ou seja:

$$\frac{\left(\bar{x} + 1{,}96\frac{3{,}2}{\sqrt{n}}\right) - \left(\bar{x} - 1{,}96\frac{3{,}2}{\sqrt{n}}\right)}{2} =$$

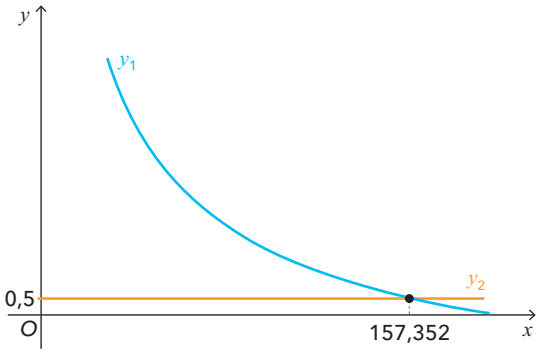
$$= \frac{2 \times 1{,}96\frac{3{,}2}{\sqrt{n}}}{2} = 1{,}96\frac{3{,}2}{\sqrt{n}} = \frac{6{,}272}{\sqrt{n}}$$

Assim, pretende-se determinar o valor de n tal que $\frac{6{,}272}{\sqrt{n}} = 0{,}5$.

Recorrendo à calculadora gráfica, e tomando x como variável independente:

$y_1 = \frac{6{,}272}{\sqrt{n}}$

$y_2 = 0{,}5$



Coordenadas do ponto de interseção: (157,352; 0,5)

A amostra deverá conter 158 semanas.

8.

a) A população são todos os utilizadores portugueses daquela plataforma de *streaming* e a amostra são os 1200 utilizadores observados.

b) Total de utilizadores: $n = 1200$

Número de utilizadores que terminaram no máximo 5 filmes: 360 + 420 = 780

Assim, a proporção amostral é:

$$\hat{p} = \frac{780}{1200} = 0{,}65$$

Nível de confiança de 90%, logo $z \approx 1{,}645$.

Intervalo de confiança:

$$\left[\hat{p} - z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\,;\hat{p} + z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$$

Substituindo os valores, obtemos:

$$\left[0{,}65 - 1{,}645\sqrt{\frac{0{,}65(1-0{,}65)}{1200}}\,;\right.$$

$$\left.0{,}65 + 1{,}645\sqrt{\frac{0{,}65(1-0{,}65)}{1200}}\right] = [0{,}6274;\ 0{,}6726]$$

Arredondando os limites do intervalo às centésimas, temos o intervalo de confiança pretendido: [0,63; 0,67]

9.

a) Sabemos que:
- $n = 64$
- $\bar{x} = 28{,}5$
- $s = 5$
- para um nível de confiança de 99%, $z \approx 2{,}576$.

Assim, substituindo os valores no intervalo $\left[\bar{x} - z\frac{s}{\sqrt{n}}\,;\,\bar{x} + z\frac{s}{\sqrt{n}}\right]$, obtemos:

$$\left[28{,}5 - 2{,}576\frac{5}{\sqrt{64}}\,;\,28{,}5 + 2{,}576\frac{5}{\sqrt{64}}\right] =$$

$$= [26{,}89;\ 30{,}11]$$

O intervalo pedido é, então, [26,89; 30,11].

$$\text{Margem de erro} = \frac{30{,}11 - 26{,}89}{2} = 1{,}61$$

b) Como o reitor afirmou que, em média, os clientes da cafetaria universitária passam 30 minutos no espaço, e 30 pertence ao intervalo [26,89; 30,11], o Joaquim não tem razão para questionar a afirmação do reitor.

c) Para diminuir a margem de erro, mantendo o nível de confiança, o Joaquim poderia aumentar o tamanho da amostra.

10. Um intervalo de 99% de confiança para o valor médio (usando o desvio-padrão amostral s) é $\left[\bar{x} - 2{,}576\frac{s}{\sqrt{n}}\,;\,\bar{x} + 2{,}576\frac{s}{\sqrt{n}}\right]$, já que $z \approx 2{,}576$.

A amplitude do intervalo é:

$$\left(\bar{x} + 2{,}576\frac{10}{\sqrt{n}}\right) - \left(\bar{x} - 2{,}576\frac{10}{\sqrt{n}}\right) =$$

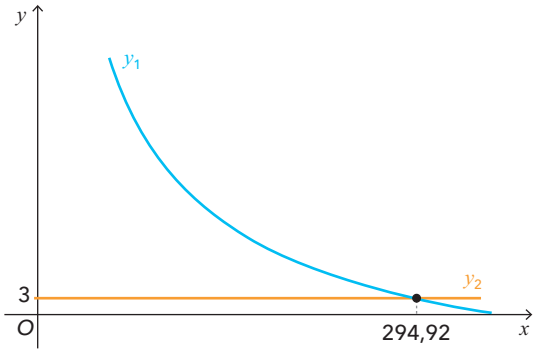
$$= 2 \times 2{,}576\frac{10}{\sqrt{n}} = \frac{51{,}52}{\sqrt{n}}$$

Assim, pretende-se determinar o valor de n tal que $\frac{51{,}52}{\sqrt{n}} = 3$.

Recorrendo à calculadora gráfica, e tomando x como variável independente:

$y_1 = \frac{51{,}52}{\sqrt{n}}$

$y_2 = 3$



Coordenadas do ponto de interseção: (294,92; 3)
A amostra deverá conter 295 clientes.

11. $n = 175 + 275 + 150 + 100 = 700$

Proporção amostral: $\hat{p} = \frac{175}{700} = 0,25$

$$\left[0,25 - z\sqrt{\frac{0,25(1-0,25)}{700}};\ 0,25 + z\sqrt{\frac{0,25(1-0,25)}{700}}\right] =$$

$= [0,25 - 0,016366z;\ 0,25 + 0,016366z]$

Amplitude $= 0,25 + 0,016366z - (0,25 - 0,016366z) =$
$= 2 \times 0,016366z$

Como se pretende que a amplitude seja 0,0641, vem que $2 \times 0,016366z = 0,0641$.

Como $2 \times 0,016366z = 0,0641 \Leftrightarrow z = \frac{0,0641}{2 \times 0,016366}$.

Assim, $z \approx 1,96$.

Sabemos, ainda, que o nível de confiança associado ao valor $z \approx 1,96$, é aproximadamente, 95%.

Exercícios–Tipo Exame – páginas 51 a 53

1. Sabemos que:
- $n = 200$
- $\bar{x} = 1,8$
- $s = 0,5$
- para um nível de confiança de 99%, $z \approx 2,576$.

Aplicando a fórmula:

$$\left[\bar{x} - z\frac{s}{\sqrt{n}};\ \bar{x} + z\frac{s}{\sqrt{n}}\right]$$

e substituindo os valores, obtemos:

$$= \left[1,8 - 2,576\frac{0,5}{\sqrt{200}};\ 1,8 + 2,576\frac{0,5}{\sqrt{200}}\right] =$$

Conservando três casas decimais nos cálculos intermédios, obtemos:

$[1,8 - 0,091;\ 1,8 + 0,091] = [1,709;\ 1,891]$

O intervalo pedido, arredondando os extremos às centésimas, é [1,71; 1,89].

O intervalo de confiança a 99% para o tempo médio de permanência é [1,71; 1,89] horas.

Como 2 horas não pertence ao intervalo de confiança, a afirmação em média, cerca de 2 horas, é incompatível com os dados (a 99%). Logo, a Amélia tem razão para duvidar da afirmação do responsável, ao nível de confiança considerado.

2. $n = 1500$

$\hat{p} = \frac{360 + 540}{1500} = \frac{900}{1500} = 0,6$

Para um nível de confiança de 95%, considera-se $z \approx 1,96$.

Aplicando a fórmula, e substituindo os valores, obtemos:

$$\left[\hat{p} - z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}};\ \hat{p} + z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right] =$$

$$= \left[0,6 - 1,96\sqrt{\frac{0,6 \times 0,4}{1500}};\ 0,6 + 1,96\sqrt{\frac{0,6 \times 0,4}{1500}}\right]$$

Conservando quatro casas decimais, nos cálculos intermédios, obtemos:

$[0,6 - 1,96 \times 0,0126;\ 0,6 + 1,96 \times 0,0126] =$

$= [0,6 - 0,0247;\ 0,6 + 0,0247] = [0,5753;\ 0,6247]$

O intervalo pedido, arredondando os extremos às centésimas, é [0,58; 0,62].

Este é o intervalo de confiança a 95% para a proporção de jovens adultos que utilizam redes sociais durante, no máximo, 60 minutos por dia.

3.

3.1 Opção (A)

A margem de erro do intervalo de confiança é igual a metade da amplitude do intervalo $\frac{45,4 - 42,6}{2} =$

$= \frac{2,8}{2} = 1,4$.

3.2 Dado que a amostra tem dimensão superior a 30, podemos determinar o intervalo de confiança, sabendo que:
- $n = 64$
- $\bar{x} = 44$
- $s = 8$
- para um nível de confiança de 90%, $z = 1,645$.

Aplicando a fórmula:

$$\left[\bar{x} - z\frac{s}{\sqrt{n}},\ \bar{x} + z\frac{s}{\sqrt{n}}\right]$$

e substituindo os valores, obtemos:

$$= \left[44 - 1,645\frac{8}{\sqrt{64}};\ 44 + 1,645\frac{8}{\sqrt{64}}\right] =$$

$= [44 - 1,645;\ 44 + 1,645] =$

$= [42,355;\ 45,645]$

O intervalo pedido, arredondando os extremos às décimas, é [42,4; 45,6].

4. Dado que a amostra tem dimensão superior a 30, podemos determinar o intervalo de confiança, sabendo que:
- $n = 225$
- $\bar{x} = 6$
- $s = 4,2$
- para um nível de confiança de 99%, $z = 2,576$.

A margem de erro deste intervalo, $\left[\bar{x} - z\frac{s}{\sqrt{n}};\ \bar{x} + z\frac{s}{\sqrt{n}}\right]$ é igual a:

$$\frac{\left(\bar{x} + z\frac{s}{\sqrt{n}}\right) - \left(\bar{x} - z\frac{s}{\sqrt{n}}\right)}{2} = \frac{2 \times z\frac{s}{\sqrt{n}}}{2} = z\frac{s}{\sqrt{n}}$$

Substituindo pelos respetivos valores, obtemos:

$2,576\frac{4,2}{\sqrt{225}} \approx 0,72$

5. Como $X \sim N(5{,}00;\ 0{,}32)$ e $n = 144$, então $\overline{X} \sim N\left(5;\ \frac{0{,}32}{\sqrt{144}}\right)$.

I – b) $\sigma_{\overline{X}} = \frac{0{,}32}{\sqrt{144}} = \frac{0{,}32}{12} = 0{,}0267 \approx 0{,}027$

II – c) Como 5,01 é maior do que 5,00, então $P(\overline{X} > 5{,}01)$ tem de ser menor do que 0,5.
Entre as opções, a única possível é 0,35.

III – a) Pela simetria da normal em torno de 5,00:
$P(\overline{X} > 5{,}02) = P(\overline{X} < 5{,}00 - 0{,}02) = P(\overline{X} < 4{,}98)$

IV – c) $P(5 < \overline{X} < 5 + 2\sigma_{\overline{X}}) = P(0 < Z < 2) \approx 0{,}47725$
A resposta correta é: I – b); II – c); III – a); IV – c)

6. Opção (B)
Tamanho da amostra: $n = 196$
Média amostral: $\overline{x} = 42$
O intervalo de confiança a 95% é $\left[42 - 1{,}96 \times \frac{\sigma}{\sqrt{196}},\ 42 + 1{,}96 \times \frac{\sigma}{\sqrt{196}}\right]$, que é igual a $]40{,}5;\ 43{,}5[$.
Assim, por exemplo:

$$42 - 1{,}96 \times \frac{\sigma}{\sqrt{196}} = 40{,}5 \Leftrightarrow 1{,}96 \times \frac{\sigma}{14} = 1{,}5 \Leftrightarrow \sigma = \frac{1{,}5 \times 14}{1{,}96}$$

Logo, $\sigma \approx 10{,}71$.

Teste Final – páginas 54 a 57

1. Um intervalo de 99% de confiança para o valor médio (usando o desvio-padrão amostral s) é:

$$\left[\overline{x} - 2{,}576\frac{s}{\sqrt{n}};\ \overline{x} + 2{,}576\frac{s}{\sqrt{n}}\right]$$

A amplitude do intervalo é:

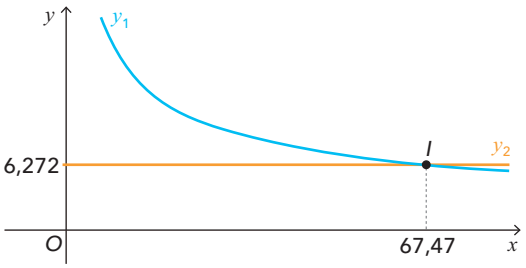
$$A = \left(\overline{x} + 2{,}576\frac{s}{\sqrt{n}}\right) - \left(\overline{x} - 2{,}576\frac{s}{\sqrt{n}}\right) = 2 \times 2{,}576\frac{s}{\sqrt{n}}$$

Como $A = 6{,}272$ e $s = 10$, obtemos a condição:

$$2 \times 2{,}576\frac{10}{\sqrt{n}} = 6{,}272$$

Para resolvermos a equação $2 \times 2{,}576\frac{10}{\sqrt{n}} = 6{,}272$, vamos recorrer às capacidades da calculadora gráfica, utilizando x como variável independente.

$f(x) = 2 \times 2{,}576\frac{10}{\sqrt{n}}$
$g(x) = 6{,}272$



$I(67{,}47;\ 6{,}272)$
$x = 67{,}47$

Como a dimensão da amostra tem de ser um número inteiro, conclui-se que a amostra terá tido, aproximadamente, 67 utentes.

2.

2.1 Como há 4 alunos, e a seleção é com reposição, existem $4^2 = 16$ amostras possíveis (considerando a ordem).

Amostra	Média
(J, J) = (6, 6)	6
(J, L) = (6, 8)	7
(J, M) = (6, 10)	8
(J, P) = (6, 12)	9
(L, J) = (8, 6)	7
(L, L) = (8, 8)	8
(L, M) = (8, 10)	9
(L, P) = (8, 12)	10
(M, J) = (10, 6)	8
(M, L) = (10, 8)	9
(M, M) = (10, 10)	10
(M, P) = (10, 12)	11
(P, J) = (12, 6)	9
(P, L) = (12, 8)	10
(P, M) = (12, 10)	11
(P, P) = (12, 12)	12

2.2 Cada amostra tem probabilidade $\frac{1}{16}$. Contando quantas vezes aparece cada valor da média:

$\overline{X}$	Frequência	Probabilidade
6	1	$\frac{1}{16}$
7	2	$\frac{2}{16}$
8	3	$\frac{3}{16}$
9	4	$\frac{4}{16}$
10	3	$\frac{3}{16}$
11	2	$\frac{2}{16}$
12	1	$\frac{1}{16}$

$$\mu_{\overline{X}} = \frac{6 \times 1 + 7 \times 2 + 8 \times 3 + 9 \times 4 + 10 \times 3 + 11 \times 2 + 12 \times 1}{16} = \frac{144}{16} = 9$$

Logo, o valor médio da distribuição de amostragem da média para as amostras do item anterior é 9.

2.3 Valor médio:

$$\mu = \frac{6 + 8 + 10 + 12}{4} = \frac{36}{4} = 9$$

Comparando:
- valor esperado da média amostral: 9
- valor médio da população: 9

A média amostral $\overline{X}$ é um estimador não enviesado (centrado) da média populacional μ, porque $\mu_{\overline{X}} = \mu$.

3. **I – b)** Os 1500 utilizadores registados são a população do estudo.

II – b) A estimativa pontual do número médio de dias/semana para todos os utilizadores é a média amostral 2,8.

III – a) Proporção que prefere estudo individual = $= \frac{90}{300} = 0{,}30$

IV – c) Se a amostra fosse 450, em vez de 300 (mantendo seleção aleatória), seria mais representativa (tende a reduzir o erro amostral).

A resposta correta é: I – b); II – b); III – a); IV – c)

4.

4.1 Temos $\mu = 50$, $\sigma = 6$ e $n = 36$.

Pelo teorema do limite central, a média amostral $\overline{X}$ tem distribuição, aproximadamente, normal com $\mu_{\overline{X}} = \mu = 50$, $\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{6}{6} = 1$.

Logo, $\overline{X} \sim N(50, 1)$.

4.2 O intervalo [48, 52] está:

- duas unidades abaixo de 50 e duas acima de 50;
- como $\sigma_{\overline{X}} = 1$, isso corresponde a ±2 desvios-padrão em torno da média.

ou seja, $P(48 \leq \overline{X} \leq 52) = P(\mu_{\overline{X}} - 2\,\sigma_{\overline{X}} \leq \overline{X} \leq \mu_{\overline{X}} + 2\,\sigma_{\overline{X}})$.

Aproximadamente 95,45% dos valores estão entre $\mu \pm 2\sigma$.

Então, $P(48 \leq \overline{X} \leq 52) \approx 0{,}9545$.

5. Atendendo a que a amostra tem dimensão superior a 30, podemos determinar o intervalo de confiança, sabendo:

- dimensão da amostra: $n = 250$
- número de visitantes que conheceram a exposição pelas redes sociais: $x = 72$
- proporção amostral: $\hat{p} = \frac{72}{250} = 0{,}288$
- para um nível de confiança de 99%, considera-se $z = 2{,}576$.

Extremos do intervalo de confiança a 99%:

$$\hat{p} \pm z \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} =$$

$$= 0{,}28800 \pm 2{,}5758 \sqrt{\frac{0{,}28800(1 - 0{,}28800)}{250}} =$$

$$= 0{,}28800 \pm 2{,}5758\sqrt{0{,}00082}$$

$$= 0{,}28800 \pm 2{,}5758 \times 0{,}02864$$

$$= 0{,}28800 \pm 0{,}07377$$

Logo:

Extremo inferior: 0,21423

Extremo superior: 0,36177

Arredondando às centésimas: [0,21; 0,36]

Como $0{,}30 \in [0{,}21; 0{,}36]$, não há razão, ao nível de confiança de 99%, para duvidar da afirmação dos 30%.

6. Opção (B)

- Dimensão da amostra: $n = 120$
- Intervalo de confiança a 95% para a proporção dos visitantes que permanecem no máximo dois dias: [0,41; 0,59]

Num intervalo de confiança do tipo $\left[\hat{p} - z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}; \hat{p} + z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right]$, os extremos são

$\underbrace{\hat{p} - z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}_{0{,}41}$; $\underbrace{\hat{p} + z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}}_{0{,}59}$

equidistantes de $\hat{p}$.

Logo, a proporção amostral é o ponto médio do intervalo:

$$\hat{p} = \frac{0{,}41 + 0{,}59}{2} = 0{,}50$$

$$\hat{p} = 0{,}50 \Leftrightarrow \frac{25 + x}{120} = 0{,}5 \Leftrightarrow 25 + x = 50 \Leftrightarrow x = 35$$

$$25 + 35 + y + 30 = 120 \Leftrightarrow y = 120 - 90 \Leftrightarrow y = 30$$

O número de visitantes que afirmaram permanecer exatamente três dias na cidade é 30.

7. Opção (C)

A proporção amostral dos que preferem a medida C é $\hat{p} = \frac{324}{900} = 0{,}36$.

A amplitude de um intervalo de confiança para uma proporção é $2z\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.

Logo:

$$0{,}06272 = 2z\sqrt{\frac{0{,}36 \times 0{,}64}{900}} \Leftrightarrow 0{,}06272 = 2z\sqrt{\frac{0{,}2304}{900}}$$

$$\Leftrightarrow 0{,}06272 = 2z\sqrt{0{,}000256} \Leftrightarrow 0{,}06272 = 2z \times 0{,}016$$

$$\Leftrightarrow 0{,}06272 = 0{,}032z$$

Assim, $z = \frac{0{,}06272}{0{,}032} = 1{,}96$.

Como $z = 1{,}96$ corresponde a 95%, o nível de confiança é 95%.

8. Opção (A)

- Proporção populacional: $p = 0{,}05$
- Dimensão da amostra (produção diária): $n = 400 \geq 30$

$np = 400 \times 0{,}05 = 20 \geq 5$

$n(1 - p) = 400 \times 0{,}95 = 380 \geq 5$

Como a dimensão da amostra é suficientemente grande ($n \geq 30$), a aproximação normal é válida.

Distribuição de amostragem de $\hat{P}$:

$$\hat{P} \sim N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$$

Desvio-padrão:

$$\sigma_p = \sqrt{\frac{0{,}05 \times 0{,}95}{400}} = \sqrt{0{,}00011875} \approx 0{,}0109 \approx 0{,}01$$

A distribuição normal que aproxima a distribuição de amostragem de $\hat{P}$ é $N(0{,}05; 0{,}01)$.