

III. Introdução à inferência estatística

Capítulo 4 Introdução à inferência estatística

Exercícios de aplicação

pág. 143

1.1 O professor pode, por exemplo, atribuir um número de 1 a 1346 a cada aluno e usar um processo aleatório (como, por exemplo, uma tabela de números aleatórios ou um sorteio) para escolher 65 números diferentes entre 1 e 1346. De seguida, seleciona os alunos correspondentes aos números sorteados.

1.2 Podemos dizer que é um método probabilístico porque todos os alunos têm a mesma probabilidade de serem escolhidos. Algumas vantagens são a simplicidade na sua aplicação e a garantia de que a amostra é representativa da população (se o sorteio for realmente aleatório).

2.1 O gerente poderá dividir o número total de clientes que passaram, nesse dia, nas caixas pela dimensão da amostra que pretende, para obter o intervalo de seleção:

$$\frac{900}{150} = 6$$

De seguida, escolhe, aleatoriamente, um número inicial entre 1 e 10 (por exemplo, 7). Por fim, seleciona o cliente com o número 7, e, sucessivamente, o $7 + 6 = 13$, o $13 + 6 = 19$, e assim por diante, até ter os 150 clientes para a sua amostra.

2.2 Este método é mais prático e rápido, especialmente quando a lista de clientes está ordenada e garante uma distribuição regular dos clientes selecionados para a amostra (pode evitar a concentração de casos semelhantes).

3.1 A população é constituída pelas 1486 pessoas que estão na sala.

3.2 A amostra será constituída pelas 12 pessoas que vão ser selecionadas de forma sistemática.

3.3 Se o primeiro sorteado é o número 16, haverá ainda $1486 - 16 = 1470$ pessoas suscetíveis de serem selecionadas. Como $1470 : 12 = 122,5 \approx 123$, haverá 123 senhas de intervalo entre dois selecionados consecutivos.

$$\begin{aligned} 16 + 123 &= 139; 139 + 123 = 262; 262 + 123 = 385; 385 + 123 = 508; 508 + 123 = 631; \\ 631 + 123 &= 754; 754 + 123 = 877; 877 + 123 = 1000; 1000 + 123 = 1123; \\ 1123 + 123 &= 1246; 1246 + 123 = 1369 \end{aligned}$$

Assim, os felizes contemplados serão os espectadores que tiverem uma das senhas pertencentes ao conjunto:

$$\{16, 139, 262, 385, 508, 631, 754, 877, 1000, 1123, 1246, 1369\}$$

4.1

	Número total de alunos	Porcentagem		Número de alunos	
		♀	♂	♀	♂
1.º Ciclo	860	55	45	473	387
2.º Ciclo	580	50	50	290	290
3.º Ciclo	1230	60	40	738	492
Secundário	1850	36	64	666	1184
Total	4520			2167	2353

4.2 $4520 \times 0,10 = 452$ alunos que deverão fazer parte da amostra.

$$1.^\circ \text{ Ciclo: } \begin{cases} 473 \times 0,10 = 47,3 \approx 47 \text{ raparigas} \\ 387 \times 0,10 = 38,7 \approx 39 \text{ rapazes} \end{cases}$$

$47 + 39 = 86$ alunos, que corresponde a 10% do total de alunos do 1.º Ciclo.

$$2.^\circ \text{ Ciclo: } \begin{cases} 290 \times 0,10 = 29 \text{ raparigas} \\ 290 \times 0,10 = 29 \text{ rapazes} \end{cases}$$

$29 + 29 = 58$ alunos, que corresponde a 10% do total de alunos do 2.º Ciclo.

$$3.^\circ \text{ Ciclo: } \begin{cases} 738 \times 0,10 = 73,8 \approx 74 \text{ raparigas} \\ 492 \times 0,10 = 49,2 \approx 49 \text{ rapazes} \end{cases}$$

$74 + 49 = 123$ alunos, que corresponde a 10% do total de alunos do 3.º Ciclo.

$$\text{Ensino Secundário: } \begin{cases} 666 \times 0,10 = 66,6 \approx 67 \text{ raparigas} \\ 1184 \times 0,10 = 118,4 \approx 118 \text{ rapazes} \end{cases}$$

$67 + 118 = 185$ alunos, que corresponde a 10% do total de alunos do Ensino Secundário.

Então, deverão ser selecionados 86 alunos do 1.º Ciclo, 58 do 2.º Ciclo, 123 do 3.º Ciclo e 185 do Ensino Secundário, num total de 452 alunos.

5. Como 20% dos alunos do agrupamento corresponde a $4365 \times 0,20 = 873$, este é o número total de alunos que deverão fazer parte da amostra.

Cálculo da proporção de alunos de cada ciclo:

$$1.^\circ \text{ Ciclo: } \frac{937}{4365} \approx 0,2147$$

$$2.^\circ \text{ Ciclo: } \frac{598}{4365} \approx 0,1370$$

$$3.^\circ \text{ Ciclo: } \frac{1236}{4365} \approx 0,2832$$

$$\text{Ensino Secundário: } \frac{1594}{4365} \approx 0,3652$$

Cálculo do número de alunos a selecionar de cada ciclo:

$$1.^\circ \text{ Ciclo: } 873 \times 0,2147 \approx 187 \text{ alunos}$$

2.º Ciclo: $873 \times 0,1370 \approx 120$ alunos

3.º Ciclo: $873 \times 0,2832 \approx 247$ alunos

Ensino Secundário: $873 \times 0,3652 \approx 319$ alunos

Assim, deverão fazer parte desta amostra 187 alunos do 1.º Ciclo, 120 do 2.º Ciclo, 247 do 3.º Ciclo e 319 do Ensino Secundário, num total de 873 alunos.

pág. 144

6.1 A população é o conjunto total de alunos do Ensino Secundário da escola, ou seja, 900 alunos.

6.2 A amostra é o conjunto de alunos inquiridos, isto é, os 120 alunos que foram escolhidos, aleatoriamente, para responder ao inquérito.

6.3 O parâmetro é a percentagem real, ou a proporção, dos 900 alunos do Ensino Secundário que preferem ler romances.

6.4 A estatística é a percentagem da amostra que prefere romances. Dos 120 alunos inquiridos, 42 referiram que preferem romances, ou seja, a estatística é $\frac{42}{120} \times 100 = 35\%$. Este valor é uma estimativa do parâmetro real para o total de alunos da escola (a população).

7. (B)

$$\hat{p} = \frac{500-153}{500} = 0,694$$

8.1 Parâmetro: proporção de alunos da escola que foram ao cinema, pelo menos, uma vez no último mês.

8.2 Estatística: $\hat{p} = \frac{83}{130} \approx 0,638$

8.3 $1 - 0,638 = 0,362 \rightarrow 36,2\%$

9.1 e 9.2 Podemos recorrer à calculadora gráfica, no modo estatístico, para obter os valores solicitados:

	List 1	List 2	List 3	List 4
SUB				
7	4	1		
8	7	1		
9	5	1		
10	6	1		

```
1Var XList :List1
1Var Freq :List2
2Var XList :List1
2Var YList :List2
2Var Freq :1
```

```
1-Variable
x̄ =5.8
Σx =58
Σx² =352
σx =1.24899959
sx =1.31656117
n =10
```

A estimativa pontual da média é 5,8 horas e a estimativa pontual do desvio padrão é aproximadamente 1,317 horas.

10.1 São $5^2 = 25$ amostras .

10.2

Amostras	$\bar{X}$	Amostras	$\bar{X}$	Amostras	$\bar{X}$	Amostras	$\bar{X}$	Amostras	$\bar{X}$
(1, 1)	1	(3, 1)	2	(6, 1)	3,5	(8, 1)	4,5	(9, 1)	5
(1, 3)	2	(3, 3)	3	(6, 3)	4,5	(8, 3)	5,5	(9, 3)	6
(1, 6)	3,5	(3, 6)	4,5	(6, 6)	6	(8, 6)	7	(9, 6)	7,5
(1, 8)	4,5	(3, 8)	5,5	(6, 8)	7	(8, 8)	8	(9, 8)	8,5
(1, 9)	5	(3, 9)	6	(6, 9)	7,5	(8, 9)	8,5	(9, 9)	9

10.3

$\bar{X}$	1	2	3	3,5	4,5	5	5,5	6	7	7,5	8	8,5	9
p_i	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{4}{25}$	$\frac{2}{25}$	$\frac{2}{25}$	$\frac{3}{25}$	$\frac{2}{25}$	$\frac{2}{25}$	$\frac{1}{25}$	$\frac{2}{25}$	$\frac{1}{25}$

10.4 $\mu = \frac{1+3+6+8+9}{5} = 5,4$

$$E(\bar{X}) = 1 \times \frac{1}{25} + 2 \times \frac{2}{25} + 3 \times \frac{1}{25} + \dots + 8,5 \times \frac{2}{25} + 9 \times \frac{1}{25} = 5,4$$

$\mu = E(\bar{X})$, logo, o estimador é não enviesado (é cêntrico).

10.5 $\sigma = 3,007$ e $\sigma_{\bar{X}} \approx 2,126$ (valores obtidos com a calculadora, modo estatístico)

10.6 $\frac{\sigma}{\sqrt{2}} \approx \frac{3,007}{\sqrt{2}} \approx 2,126$

O valor obtido é igual ao desvio padrão da distribuição de amostragem da média.

11. Uma vez que são $5^3 = 125$ amostras , sugere-se a divisão da turma em cinco grupos e cada grupo faz o estudo de 25 amostras: determinar as amostras, a média de cada uma e os diferentes valores que esta pode assumir. Por exemplo, o grupo 1 estuda as amostras que começam com o elemento 1, o grupo 2, as amostras que começam pelo elemento 3, e assim sucessivamente. Quando se juntam os resultados obtidos por todos os grupos, dever-se-á obter:

$\bar{X}$	1	1,67	2,33	2,67	3	3,33	3,67	4	4,33	4,67	5	5,33
p_i	$\frac{1}{125}$	$\frac{3}{125}$	$\frac{3}{125}$	$\frac{3}{125}$	$\frac{1}{125}$	$\frac{9}{125}$	$\frac{3}{125}$	$\frac{9}{125}$	$\frac{9}{125}$	$\frac{3}{125}$	$\frac{12}{125}$	$\frac{6}{125}$
$\bar{X}$	5,67	6	6,33	6,67	7	7,33	7,67	8	8,33	8,67	9	
p_i	$\frac{9}{125}$	$\frac{13}{125}$	$\frac{6}{125}$	$\frac{9}{125}$	$\frac{6}{125}$	$\frac{3}{125}$	$\frac{6}{125}$	$\frac{4}{125}$	$\frac{3}{125}$	$\frac{3}{125}$	$\frac{1}{125}$	

$$\mu = 5,4 = E(\bar{X}) ; \sigma_{\bar{X}} \approx 1,736$$

12.1 (B)

12.2 (C)

$$\sigma_{\bar{X}} = \frac{9,25}{\sqrt{40}} \approx 1,463 \text{ g}$$

$$13.1 E(\bar{X}) = 17,5 \text{ kg}$$

$$13.2 \sigma_{\bar{X}} = \frac{2,75}{\sqrt{35}} \approx 0,465 \text{ kg}$$

$$14.1 \mu_{\bar{X}} = 1250 \text{ horas}$$

$$14.2 \sigma_{\bar{X}} = \frac{110}{\sqrt{45}} \approx 16,398 \text{ horas}$$

$$14.3.1 U = \frac{\bar{X} - 1250}{16,398} \Leftrightarrow \bar{X} = 16,398U + 1250$$

$$\begin{aligned} P(1200 < \bar{X} < 1280) &= P(1200 - 1250 < 16,398U < 1280 - 1250) \approx \\ &\approx P(-3,05 < U < 1,83) = \Phi(1,83) - 1 + \Phi(3,05) \approx 0,9664 - 1 + 0,9989 = \\ &= 0,9653 \rightarrow 96,53\% \end{aligned}$$

(tabela da página 62 do Manual)

$$\begin{aligned} 14.3.2 P(|\bar{X} - \mu| < 0,4) &= P\left(|U| < \frac{0,4}{16,398}\right) \approx P(|U| < 0,02) = 2\Phi(0,02) - 1 = 2 \times 0,5080 - 1 = \\ &= 0,016 \rightarrow 1,6\% \end{aligned}$$

15. O intervalo de 90% para μ é:

$$\left] \bar{x} - 1,645 \times \frac{\sigma}{\sqrt{n}}; \bar{x} + 1,645 \times \frac{\sigma}{\sqrt{n}} \right[$$

Neste caso, tem-se:

$$\left] 150 - 1,645 \times \frac{50}{\sqrt{100}}; 150 + 1,645 \times \frac{50}{\sqrt{100}} \right[=]141,775; 158,225[$$

16. $n = 80$, $\bar{x} = 127$ e $s = 9$ O intervalo de confiança de 90% para μ é:

$$\left] 127 - 1,645 \times \frac{9}{\sqrt{80}}; 127 + 1,645 \times \frac{9}{\sqrt{80}} \right[=]125,34; 128,66[$$

17. O intervalo de 95% para μ é:

$$\left] \bar{x} - 1,96 \times \frac{\sigma}{\sqrt{n}}; \bar{x} + 1,96 \times \frac{\sigma}{\sqrt{n}} \right[=]\bar{x} - 1,96 \times 0,5; \bar{x} + 1,96 \times 0,5[=]\bar{x} - 0,98; \bar{x} + 0,98[$$

$P(\bar{x} - 0,98 \leq \mu \leq \bar{x} + 0,98) = 0,95$, o que significa que a probabilidade de o intervalo $] \bar{x} - 0,98; \bar{x} + 0,98[$ conter o valor médio μ é 0,95.

18. $n = 80$; $\bar{x} = 1,6$ e $s = 0,05$ O intervalo de 95% para μ é:

$$\left] 1,6 - 1,96 \times \frac{0,05}{\sqrt{80}}; 1,6 + 1,96 \times \frac{0,05}{\sqrt{80}} \right[=]1,59; 1,61[$$

19.1 $n = 80$, $\bar{x} = 70$ e $s = 4$

Um intervalo de confiança de 90% para μ é:

$$\left] 70 - 1,645 \times \frac{4}{\sqrt{80}}; 70 + 1,645 \times \frac{4}{\sqrt{80}} \right[=]69,26; 70,74[$$

19.2 $n = 100$, $\bar{x} = 70$ e $s = 4$

Um intervalo de confiança de 95% para μ é:

$$\left] 70 - 1,96 \times \frac{4}{\sqrt{100}}; 70 + 1,96 \times \frac{4}{\sqrt{100}} \right[=]69,216; 70,784[$$

19.3 $n = 80$, $\bar{x} = 60$ e $s = 4$

Um intervalo de confiança de 99% para μ é:

$$\left] 60 - 2,576 \times \frac{4}{\sqrt{80}}; 60 + 2,576 \times \frac{4}{\sqrt{80}} \right[=]58,85; 61,15[$$

20.1 $\sigma = 10$, $n = 50$ e $\bar{x} = 994$

O intervalo de confiança de 90% para μ é:

$$\left] 994 - 1,645 \times \frac{10}{\sqrt{50}}; 994 + 1,645 \times \frac{10}{\sqrt{50}} \right[=]991,67; 996,33[$$

O intervalo de confiança de 99% para μ é:

$$\left] 994 - 2,576 \times \frac{10}{\sqrt{50}}; 994 + 2,576 \times \frac{10}{\sqrt{50}} \right[=]990,36; 997,64[$$

À medida que o grau de confiança aumenta, a amplitude do intervalo também aumenta.

20.2 $n = 200$ e $\bar{x} = 994$

O intervalo de confiança de 90% para μ é:

$$\left] 994 - 1,645 \times \frac{10}{\sqrt{200}}; 994 + 1,645 \times \frac{10}{\sqrt{200}} \right[=]992,84; 995,16[$$

20.3 À medida que a dimensão da amostra aumenta, a amplitude do intervalo diminui.

21. Sabe-se que para $s = 29$ e $n = 40$, o intervalo de confiança de 95% para μ é $]160,178[$.

Sabe-se ainda que:

$$\bar{x} = \frac{178 + 160}{2} = 169$$

Assim, o intervalo de confiança de 99%, para a mesma amostra, resulta dos seguintes valores:

$$\bar{x} = 994, s = 29 \text{ e } n = 40$$

Logo:

$$\left] 169 - 2,576 \times \frac{29}{\sqrt{40}}; 169 + 2,576 \times \frac{29}{\sqrt{40}} \right[=]157,188; 180,812[$$

Arredondando às unidades, obtemos $]157,181[$.

22. O intervalo é $]32,38[$, o nível de confiança é 90% e o tamanho da amostra é 150.

A média amostral é o ponto médio do intervalo, ou seja $\bar{x} = \frac{32+38}{2} = 35$.

A margem de erro é $\varepsilon = 38 - 35 = 3$.

$$\varepsilon = 1,645 \times \frac{s}{\sqrt{150}} \Leftrightarrow 3 = 1,645 \times \frac{s}{\sqrt{150}} \Leftrightarrow s \approx 22,343$$

O desvio padrão amostral, arredondado às unidades, é 22.

23. $n = 50$, $\bar{x} = 4$ e $s = 1,2$

$$\varepsilon = z \times \frac{s}{\sqrt{n}} \text{ e, para 95\% de confiança, } z = 1,960, \text{ logo } \varepsilon = 1,960 \times \frac{1,2}{\sqrt{50}} \Leftrightarrow \varepsilon \approx 0,333$$

24. A proporção amostral dá-nos a frequência relativa dos eleitores a favor do candidato A.

Neste caso, $\hat{p} = \frac{350}{500} = 0,7$, ou seja, 70% dos eleitores da amostra são a favor do candidato A.

25. $n = 50$

$$\hat{p} = \frac{40}{50} = 0,8 = 80\%$$

pág. 147

26. *Trabalho de pesquisa.*

27. Como $n = 400$, podemos considerar que é uma distribuição aproximadamente normal com valor médio 0,3 e desvio-padrão:

$$\sigma = \sqrt{\frac{0,3 \times (1-0,3)}{400}} \approx 0,02$$

28. $n = 500$

$$\bar{x} = \frac{300}{500} = 0,6$$

$$\sigma = \sqrt{\frac{0,6 \times (1-0,6)}{500}} \approx 0,022$$

29. *Relatório.*

30.1 $n = 1500$ e $\hat{p} = 76\%$

A margem de erro é:

$$z \sqrt{\frac{p(1-p)}{n}} = 1,645 \times \sqrt{\frac{0,76 \times 0,24}{1500}} = 0,0181 = 1,8\%$$

Portanto, a margem de erro é inferior a 5%.

30.2 Margem de erro: $1,645 \times \sqrt{\frac{0,76 \times 0,24}{3000}} = 0,0128 = 1,28\% \approx 0,01$

30.3 À medida que a dimensão da amostra aumenta, a margem de erro e a amplitude do intervalo diminuem.

31. $n = 1200$ e $p = 40\%$

Intervalo de confiança de 90%:

$$\left[0,4 - 1,645 \times \sqrt{\frac{0,4 \times 0,6}{1200}}; 0,4 + 1,645 \times \sqrt{\frac{0,4 \times 0,6}{1200}} \right] =]0,377; 0,423[$$

Intervalo de confiança de 95%:

$$\left[0,4 - 1,96 \times \sqrt{\frac{0,4 \times 0,6}{1200}}; 0,4 + 1,96 \times \sqrt{\frac{0,4 \times 0,6}{1200}} \right] =]0,372; 0,428[$$

Intervalo de confiança de 99%:

$$\left[0,4 - 2,576 \times \sqrt{\frac{0,4 \times 0,6}{1200}}; 0,4 + 2,576 \times \sqrt{\frac{0,4 \times 0,6}{1200}} \right] =]0,364; 0,436[$$

À medida que o nível de confiança aumenta, a amplitude do intervalo também aumenta.

32. Esta afirmação significa que o intervalo $]77\%, 83\%[$ é um intervalo de confiança de 95% para a proporção pretendida. 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 ETAR.

33. O intervalo $]35\%, 45\%[$ é um intervalo de confiança de 95% para a proporção de pessoas que consideram que a droga é o problema mais sério que afeta a adolescência.

A margem de erro é:

$$\frac{45-35}{2} = 5\%$$

34. $\sigma = 0,5$ e $\varepsilon \leq 0,01$

Dimensão da amostra para um intervalo de 90% de confiança:

$$n = \left(\frac{1,645 \times 0,5}{0,01} \right)^2 \approx 6765,06$$

A dimensão da amostra deverá ser 6766.

Dimensão da amostra para um intervalo de 95%:

$$n = \left(\frac{1,960 \times 0,5}{0,01} \right)^2 = 9604$$

35. $\varepsilon = 0,03$ e $p = 0,45$

Dimensão da amostra para um intervalo de 95% de confiança:

$$n = \left(\frac{z}{\varepsilon} \right)^2 \times p \times (1 - p) = \left(\frac{1,96}{0,03} \right)^2 \times 0,45 \times (1 - 0,45) = 1056,44$$

A dimensão da amostra deve ser 1057.