

Ficha 3 — Intervalos de confiança para o valor médio

12.º Ano

Matemática A

4 páginas

12 exercícios

Objetivo: construir e interpretar intervalos de confiança para o valor médio de uma população, calcular a margem de erro, usar a fórmula geral para qualquer nível de confiança e dimensionar uma amostra para uma precisão pretendida.

Revisão rápida — estimativa intervalar

- Uma **estimativa pontual** (por exemplo $\bar{x}$) dá um único valor; uma **estimativa intervalar** dá um **intervalo** de valores plausíveis para o parâmetro, acompanhado de um **nível de confiança**.
- **Margem de erro** (para a média, com desvio padrão populacional σ conhecido): $MA = z \times \sigma / \sqrt{n}$.
- **Intervalo de confiança a $100(1 - \alpha)\%$** : $\bar{x} \pm z_{1-\alpha/2} \times \sigma / \sqrt{n}$, onde $z_{1-\alpha/2}$ é o quantil da normal padrão tal que $\Phi(z) = 1 - \alpha/2$.
- Valores a memorizar: $90\% \rightarrow z = 1,645$; $95\% \rightarrow z = 1,96$; $99\% \rightarrow z = 2,576$.
- **Interpretação correta:** a partir de uma amostra o intervalo pode ou não conter μ — não sabemos qual. O que o nível de confiança diz é que o **método** produz intervalos que contêm μ em 95% dos casos, quando se repetisse a amostragem muitas vezes.
- **Como estreitar o intervalo:** aumentar o tamanho da amostra n , ou diminuir o nível de confiança (abdicando de garantia).
- **Dimensão da amostra** para uma margem de erro pretendida MA : resolve-se $n \geq (z \sigma / MA)^2$ e arredonda-se **sempre por excesso**.

Exercícios

1. Numa amostra de 100 observações obteve-se $\bar{x} = 12$ e sabe-se que $\sigma = 3$. Constrói o intervalo de confiança a 95% para μ e indica a margem de erro.

2. Explica o significado da margem de erro obtida na questão 1 e o que significa dizer que o nível de confiança é 95%.

3. Escreve a fórmula geral do intervalo de confiança a $100(1 - \alpha)\%$ para o valor médio e indica o valor de z para os níveis de 90%, 95% e 99%, confirmando cada um com o valor de $\Phi(z)$.

4. Com os mesmos dados ($\bar{x} = 12$, $\sigma = 3$, $n = 100$), constrói o intervalo de confiança a 90%.

5. Constrói ainda o intervalo de confiança a 99% e compara a sua largura com as dos intervalos das questões 1 e 4, explicando a relação entre nível de confiança e largura.

6. Mantendo $\bar{x} = 12$, $\sigma = 3$ e o nível de 95%, calcula o intervalo para $n = 400$ e compara a margem de erro com a da questão 1.

7. Um investigador afirma: «recolhi uma amostra maior, por isso o meu intervalo ficou mais estreito e agora é mais fiável». Explica em que sentido esta afirmação é correta e em que sentido deve ser lida com cuidado.

8. Pretende-se um intervalo de confiança a 95% com margem de erro não superior a 0,5, sabendo que $\sigma = 3$. Determina a dimensão mínima da amostra necessária e verifica o resultado.

9. Interpretação: explica o que está errado (e o que está correto) nas afirmações seguintes: (a) «há 95% de probabilidade de μ pertencer a este intervalo»; (b) «95% dos elementos da população pertencem a este intervalo».

10. Simulação: numa simulação construíram-se 2000 intervalos de confiança a 95% ($n = 25$, $\sigma = 1$) para uma população com $\mu = 0$ e verificou-se que 1891 continham μ . Comenta este resultado: o que confirma e por que razão não é exatamente 95%?

11. (Python) O programa seguinte simula 2000 intervalos de confiança a 95% e conta quantos contêm μ . Executa-o e indica exatamente o que é impresso, comparando com o resultado da questão 10.

```
import random

random.seed(0)
mu = 0.0
sigma = 1.0
n = 25
rep = 2000
contem = 0

for _ in range(rep):
    amostra = [random.gauss(mu, sigma) for _ in range(n)]
    media = sum(amostra) / n
    margem = 1.96 * sigma / (n ** 0.5)
    if media - margem <= mu <= media + margem:
        contem += 1

print('Intervalos que contem mu:', contem)
print('Proporcao:', round(contem / rep, 4))
```

12. Aplicação: uma empresa quer estimar o tempo médio de espera dos clientes. Numa amostra de 64 clientes obteve-se $\bar{x} = 8,5$ minutos e sabe-se que $\sigma = 4$ minutos. Constrói o intervalo de confiança a 95% e explica o que a empresa deve concluir.

Soluções — Intervalos de confiança para o valor médio

1. A margem de erro é $MA = 1,96 \times 3/\sqrt{100} = 1,96 \times 0,3 = 0,588$. O intervalo a 95% é $12 \pm 0,588$, isto é **[11,412; 12,588]**.
 2. A **margem de erro** (0,588) é o afastamento máximo que admitimos entre a estimativa $\bar{x}$ e o parâmetro μ , para o nível de confiança escolhido. Dizer que o nível de confiança é 95% significa que, **se repetíssemos a amostragem muitas vezes** e construíssemos o intervalo sempre do mesmo modo, cerca de **95% desses intervalos conteriam μ** . É uma propriedade do **método**, não de um intervalo particular.
 3. A fórmula geral é $\bar{x} \pm z_{1-\alpha/2} \times \sigma/\sqrt{n}$, com $\Phi(z_{1-\alpha/2}) = 1 - \alpha/2$. Para 90%: $1 - \alpha/2 = 0,95 \Rightarrow z = 1,645$ ($\Phi(1,645) \approx 0,9500$). Para 95%: $1 - \alpha/2 = 0,975 \Rightarrow z = 1,96$ ($\Phi(1,96) = 0,9750$). Para 99%: $1 - \alpha/2 = 0,995 \Rightarrow z = 2,576$ ($\Phi(2,576) \approx 0,9950$).
 4. $MA = 1,645 \times 0,3 = 0,4935$, logo o intervalo a 90% é $12 \pm 0,4935 =$ **[11,5065; 12,4935]**.
 5. Para 99%: $MA = 2,576 \times 0,3 = 0,7728$, logo o intervalo é **[11,2272; 12,7728]**. Comparando as larguras: 90% $\rightarrow 0,987$; 95% $\rightarrow 1,176$; 99% $\rightarrow 1,5456$. Quanto **maior o nível de confiança, mais largo** é o intervalo: para garantir com mais segurança que μ está dentro dele, é preciso admitir uma gama maior de valores plausíveis. Confiança e precisão são objetivos em tensão.
 6. $MA = 1,96 \times 3/\sqrt{400} = 1,96 \times 0,15 = 0,294$, exatamente **metade** da margem da questão 1 (0,588), pois $\sqrt{400} = 2\sqrt{100}$. O intervalo é **[11,706; 12,294]** — com quatro vezes mais observações, a estimativa é duas vezes mais precisa.
 7. A afirmação é **correta no essencial**: com a mesma confiança de 95%, um n maior reduz a margem de erro ($\sigma/\sqrt{n}$ diminui) e o intervalo fica mais estreito, o que corresponde a uma estimativa mais precisa — supondo que a amostra maior foi recolhida com o **mesmo processo aleatório**. Deve, no entanto, ser lida com cuidado em dois pontos: (i) maior precisão **não corrige enviesamento** — uma amostra grande mas mal recolhida continua a dar uma estimativa errada, apenas com aparência de rigor; (ii) a precisão melhora com $\sqrt{n}$, pelo que o ganho é cada vez menor, e há sempre um custo associado a recolher mais dados.
 8. Exige-se $MA \leq 0,5$, ou seja $1,96 \times 3/\sqrt{n} \leq 0,5$. Resolvendo: $\sqrt{n} \geq 1,96 \times 3/0,5 = 11,76$, donde $n \geq 11,76^2 = 138,2976$. Como a dimensão tem de ser um número inteiro e a desigualdade é $n \geq 138,2976$, arredonda-se **por excesso: $n = 139$** . Verificação: com $n = 139$ a margem é $1,96 \times 3/\sqrt{139} \approx 0,4987 \leq 0,5$ ✓; com $n = 138$ seria $0,5005 > 0,5$ ✗ — o que mostra por que razão não se pode arredondar por defeito.
 9. (a) **Incorreta** na forma: μ é um valor **fixo** (embora desconhecido), não aleatório — depois de calculado, o intervalo ou contém μ ou não contém; a probabilidade 95% refere-se ao método (à proporção de intervalos que conteriam μ em amostragens repetidas). Uma formulação aceitável é: «o método usado produz intervalos que contêm μ em 95% dos casos». (b) **Totalmente errada**: o intervalo é um intervalo de confiança para a **média** (um parâmetro), não um intervalo que contenha 95% dos indivíduos da população. Para conter 95% dos indivíduos seria preciso usar $\mu \pm 1,96\sigma$ — um intervalo muito mais largo, e de natureza diferente.
 10. Dos 2000 intervalos construídos, **1891 continham μ** , o que corresponde a uma proporção de **0,9455** — muito próxima de 0,95. Isto confirma o significado do nível de confiança: o método produz intervalos que «apanham» μ em cerca de 95% das vezes. Não dá exatamente 95% porque uma simulação é uma **amostra finita** (2000 repetições, não infinitas) e existe flutuação própria do gerador aleatório; variando a semente obtêm-se valores como 0,9470, 0,9480 ou 0,9545, todos em torno de 0,95.
 11. O programa imprime **Intervalos que contem mu: 1891 e Proporcao: 0.9455**. Reproduce, portanto, exatamente o resultado descrito na questão 10 — o que é natural, pois o código usa a mesma semente (`random.seed(0)`), garantindo uma sequência de números pseudoaleatórios reprodutível. Alterando a semente, o número de intervalos que contêm μ varia ligeiramente (por exemplo 1894 ou 1896 em 2000), mas mantém-se próximo de 95%: é precisamente essa estabilidade em torno de 0,95 que dá sentido à interpretação do nível de confiança.
 12. A margem de erro é $MA = 1,96 \times 4/\sqrt{64} = 1,96 \times 0,5 = 0,98$ minutos, logo o intervalo de confiança a 95% é $8,5 \pm 0,98$, isto é **[7,52; 9,48]** minutos. A empresa deve concluir que o tempo médio de espera de **todos** os clientes está plausivelmente entre cerca de 7,5 e 9,5 minutos, com 95% de confiança no método usado. Deve ter em atenção que este intervalo diz respeito à média e não ao tempo de espera de um cliente individual, e que a conclusão pressupõe que a amostra de 64 clientes foi recolhida aleatoriamente.
- Nota:** a ideia mais subtil desta ficha é a da questão 9 — μ não é aleatório, o intervalo é que é. Compreender isso evita o erro mais comum de toda a inferência estatística.