

Resumo 3 — Inferência estatística: estimar, medir a incerteza, decidir

11.º Ano MACS 2 páginas Resumo

Objetivo: consolidar **parâmetro vs estimador**, a **amostra aleatória** e a variabilidade amostral, a **distribuição de amostragem** com o **TLC** e o desvio **s (n – 1)**, e os **intervalos de confiança** (média e proporção) — o Tema 3 num quadro.

Do parâmetro à estimativa

Conceito	O que é / regra
Parâmetro (grego)	valor fixo da população : μ , p , σ — é o que se quer saber
Estimador (latino)	valor calculado na amostra : $\bar{x}$, $\hat{p}$, s — é o que se tem
Amostra aleatória	sem ela, não se quantifica o erro com probabilidade (conveniência → viés , que nem n grande corrige)
Com vs sem reposição	com = escolhas independentes · sem = a urna muda → dependentes
Variabilidade amostral	tantas estimativas quantas amostras → a estimativa pontual é limitada (falta a margem)

Distribuição de amostragem, TLC e s

Ferramenta	Regra
Distribuição de amostragem	do estimador em todas as amostras (não o histograma da amostral) — variabilidade decrece com n ($\sigma/\sqrt{n}$)
TLC	para n suficientemente grande , a distribuição de $\bar{x}$ é Normal , qualquer que seja a população — dispensa enumerar ($C(1000,30) \approx 10^{57}$)
s vs s'	$s = \sqrt{(\sum(x-\bar{x})^2/(n-1))}$ vs $s' = \sqrt{(.../n)}$ — usa-se s (melhor estimativa de σ ; s' subestima)
σ de $\bar{x}$	$\sigma/\sqrt{n}$ — dobrar n só divide por $\approx 1,4$ (a raiz engole o ganho)

Estratégia do tema: definir população e parâmetro → desenhar o processo (aleatório? com/sem reposição?) → reconhecer a distribuição do estimador → invocar o TLC com as condições → calcular com **s** → passar ao **intervalo**.

Intervalos de confiança — a régua de μ

Nível	Multiplicador	Fórmula para μ
95 %	1,96	$\bar{x} \pm 1,96 \cdot \sigma / \sqrt{n}$
90 %	1,645	$\bar{x} \pm 1,645 \cdot \sigma / \sqrt{n}$
99 %	2,576	$\bar{x} \pm 2,576 \cdot \sigma / \sqrt{n}$

- σ desconhecido $\rightarrow$ substituir por s · **margem = metade da amplitude**.
- Leitura correta:** «a probabilidade de o intervalo conter μ é 95 %» — o **aleatório está nos limites** (μ é fixo); 100 amostras dariam ~95 intervalos, mas calcula-se **um** e não se sabe se é dos 95 $\rightarrow$ «confiança».
- Alavancas da largura:** confiança $\uparrow \rightarrow$ **mais largo** (z) · $n \uparrow \rightarrow$ **mais estreito** ($\pm \sqrt{n}$) · $\sigma \uparrow \rightarrow$ mais largo.
- Proporção:** p é média de 0/1 (variância $p(1-p)$) $\rightarrow \hat{p} \pm z \cdot \sqrt{(\hat{p}(1-\hat{p}))/n}$ — sondagens: «52 % $\pm$ 3 %» = [49 %; 55 %] (sem vitória declarada!) — e a margem **não cobre viés** (não-respostas).

Erros frequentes

- $\bar{x}$ = valor verdadeiro (é estimativa) · qualquer amostra serve (tem de ser aleatória).
- TLC diz que a população é normal (diz da distribuição de $\bar{x}$) · invocá-lo com $n = 3$.
- Usar s' para estimar σ (usa-se s , com $n - 1$).
- « μ tem 95 % de probabilidade de estar» (o aleatório é do intervalo).
- « n maior dá mais confiança» (dá largura menor, à mesma confiança) · usar σ da população quando é desconhecida.

Para reter: parâmetro = **população** (grego), estimador = **amostra** (latino) · aleatória $\rightarrow$ erro **mensurável**; conveniência $\rightarrow$ **viés** · «tantas estimativas quantas amostras» · distribuição de amostragem = do **estimador entre amostras** · TLC: **n grande $\rightarrow \bar{x}$ Normal, qualquer população · s com $n - 1$ (não s')** · σ de $\bar{x} = \sigma / \sqrt{n}$ · IC95 = $\bar{x} \pm 1,96 \cdot \sigma / \sqrt{n}$ (90 % $\rightarrow 1,645$; 99 % $\rightarrow 2,576$; $\sigma \rightarrow s$) · margem = **metade da amplitude** · leitura: **o intervalo contém μ** · proporção: $\hat{p} \pm z \cdot \sqrt{(\hat{p}(1-\hat{p}))/n}$ — e sempre: **estimativa + margem + nível + limitação**.

Série MACS11 — fichas 10–12 cobram em detalhe cada secção; a AT 16 (simulação do IC em Python) comprova a cobertura; os flashcards de Inferência (item 24), o teste T2 (item 26) e o simulado (item 28) complementam a revisão.