

FRENTE · 1 / 12

1

População

FRENTE · 2 / 12

2

Amostra

FRENTE · 3 / 12

3

Variável estatística

FRENTE · 4 / 12

4

Discreta vs. contínua

VERSO

1

Conjunto **total** de indivíduos/objeto de interesse (ex.: todos os alunos do 10.º ano do país) — de onde vêm os dados.

VERSO

2

Subconjunto **representativo** da população, usado por praticidade (sorteio / estratificada) — *del*a se infere para a população.

VERSO

3

Característica observada que **varia** entre indivíduos. Pode ser **qualitativa** (categórica) ou **quantitativa** (numérica).

VERSO

4

Discreta: conta — valores isolados (nº de faltas).
Contínua: mede — qualquer valor do intervalo (massa, tempo); usa classes no histograma.

FRENTE · 5 / 12

5

Histograma

FRENTE · 6 / 12

6

Mediana

FRENTE · 7 / 12

7

AIQ (amplitude interquartil)

FRENTE · 8 / 12

8

Desvio-padrão (σ)

VERSO

5

Gráfico de classes **ordenadas e contíguas** (barras coladas); a área é proporcional à frequência — ideal p/ contínuos; usam-se larguras iguais (ou frequência por amplitude).

VERSO

6

Valor **central quando ordenados** (par de meio, se n par): divide a amostra em 50/50. **Robusta** a outliers — escolha em assimetrias.

VERSO

7

AIQ = **Q3 - Q1** — largura da caixa do boxplot: abrange os 50% centrais. Mede dispersão *robusta* (ignora caudas).

VERSO

8

√ da variância: $\sigma = \sqrt{(\sum(x-\bar{x})^2/n)}$ — afastamento típico dos pontos à média; **mesma unidade** dos dados (σ^2 fica em unidade²).

FRENTE · 9 / 12

9

Outlier (regra dos 1,5·AIQ)

FRENTE · 10 / 12

10

Coefficiente de correlação r

FRENTE · 11 / 12

11

Reta de regressão

FRENTE · 12 / 12

12

Enviesamento (viés de seleção)

VERSO

9

Valor fora de $[Q1 - 1,5 \cdot AIQ ; Q3 + 1,5 \cdot AIQ]$. Investigar (erro?) e mostrar com/senão ele — distorce média e σ (não a mediana).

VERSO

10

Pearson: mede a **associação linear** — sinal = direção, $|r|$ = força, **$r = 0$ = sem relação linear** (pode haver curva).
Correlação $\neq$ causalidade.

VERSO

11

$\hat{y} = a \cdot x + b$ escolhida para **minimizar os quadrados das distâncias** aos pontos — serve para prever y (fiável só entre os dados; nada de extrapolar).

VERSO

12

Quando a amostra **não representa a população** (só online/voluntários) → conclusões perfeitas sobre a amostra errada. Evita-se com amostragem aleatória/estratificada.