

FRENTE · 1 / 12

1

População vs amostra

FRENTE · 2 / 12

2

Variável (num estudo)

FRENTE · 3 / 12

3

Fases de um procedimento estatístico

FRENTE · 4 / 12

4

Amostra representativa vs enviesada

VERSO

1

População: todo o conjunto estudado. Amostra: subconjunto observado, de onde vêm os dados

VERSO

2

Caraterística medida em cada indivíduo — varia de caso para caso (horas, notas, idades)

VERSO

3

Pergunta → recolha → organização → análise → interpretação/conclusão — a pergunta é sempre a primeira

VERSO

4

Probabilística (aleatória/sistemática/estratificada) generaliza; conveniência e voluntários enviesam — representatividade antes da dimensão

FRENTE · 5 / 12

5

Dados discretos vs contínuos

FRENTE · 6 / 12

6

Histograma — regras de construção

FRENTE · 7 / 12

7

Média vs mediana

FRENTE · 8 / 12

8

Variância amostral s^2

VERSO

5

Discreto: inteiros contáveis (n.º de irmãos). Contínuo: qualquer valor de medição (tempo, massa)

VERSO

6

Classes de amplitude igual, retângulos colados, limites [a, b) — diagrama de áreas: altura $\propto$ frequência

VERSO

7

Média: $\Sigma x_i/n$, usa tudo mas puxa-se com outliers.
Mediana: valor central ordenado — robusta (rendimentos)

VERSO

8

$s^2 = \Sigma (x_i - \bar{x})^2 \div (n - 1)$ — em amostras divide-se por $n - 1$; desvio-padrão $s = \sqrt{s^2}$

FRENTE · 9 / 12

9

Coeficiente de correlação r

FRENTE · 10 / 12

10

Reta de regressão

FRENTE · 11 / 12

11

Correlação $\neq$ causalidade

FRENTE · 12 / 12

12

As duas conferências das tabelas de frequências

VERSO

9

$-1 \leq r \leq 1$ · sinal = direção (+/-) · $|r|$ = força · $r \approx 0$ = sem associação linear (pode haver não linear)

VERSO

10

$\hat{y} = a + b \cdot x$, com $b = r \cdot (s_y/s_x)$ e $a = \bar{y} - b \cdot \bar{x}$ — serve para inferir $\hat{y}$; não extrapolar para longe dos dados

VERSO

11

Gelados e afogamentos correlacionam-se pelo calor (3.º fator) — para provar causa exige-se mecanismo ou experimentação

VERSO

12

$\sum f_i = n$ e $\sum f_r = 1$ — falha uma, tabela e gráfico estão errados