Formulário de Estatística – 10º ano

Sinal de somatório

$$\sum_{k=p}^{n} x_k \text{, com } n \geqslant p \text{ inteiros}$$

Lê-se: Somatório de x_k para k de p a n

k: índice do somatório

p: limite inferior do somatório

n: limite superior do somatório

$n-p+1$: número de parcelas

p. 181

- $\sum_{i=3}^{6} 3^i = 3^3 + 3^4 + 3^5 + 3^6$
- $x_1 y_1 + x_2 y_2 + \ldots + x_9 y_9 = \sum_{i=1}^{9} x_i y_i$

Propriedades dos somatórios

- $\sum_{k=p}^{n} (a_k + b_k) = \sum_{k=p}^{n} a_k + \sum_{k=p}^{n} b_k$, com $p \leqslant n$
- $\sum_{k=p}^{n} c a_k = c \sum_{k=p}^{n} a_k$, com $p \leqslant n$ e c constante
- $\sum_{k=p}^{n} c = c(n-p+1)$, com c constante
- $\sum_{k=1}^{n} c = nc$, com c constante
- $\sum_{k=p}^{n} a_k = \sum_{k=p}^{m} a_k + \sum_{k=m+1}^{n} a_k$, com $p \leqslant m < n$

pp. 183 a 185

- $\sum_{i=1}^{5} (2+3i) = \sum_{i=1}^{5} 2 + \sum_{i=1}^{5} 3i$
- $\sum_{i=1}^{3} 2i = 2 \sum_{i=1}^{3} i$
- $\sum_{i=1}^{10} 5i = \sum_{i=1}^{5} 5i + \sum_{i=6}^{10} 5i$
- $\sum_{i=1}^{5} x = 5x$

Média de uma amostra

- $\overline{x} = \dfrac{\sum_{i=1}^{n} x_i}{n}$ (dados simples)
- $\overline{x} = \dfrac{\sum_{j=1}^{m} \tilde{x}_j n_j}{n}$ (dados agrupados)

pp. 186 e 187

- $\underset{\sim}{x} = (1, 0, 3)$; $\overline{x} = \dfrac{\sum_{i=1}^{3} x_i}{3} = \dfrac{1+0+3}{3} = \dfrac{4}{3}$
- $\underset{\sim}{x} = (1, 1, 1, 1, 3, 3, 3, 0, 0)$

$\overline{x} = \dfrac{\sum_{j=1}^{3} \tilde{x}_j n_j}{n}$, ou seja, $\overline{x} = \dfrac{1 \times 4 + 3 \times 3 + 0 \times 2}{9} = \dfrac{13}{9}$

Propriedades da média

- $\underset{\sim}{x} = (x_1, x_2, \ldots, x_n)$

  $\underset{\sim}{y} = (ax_1 + h, ax_2 + h, \ldots, ax_n + h)$

  $\overline{y} = a\overline{x} + h$
- Sendo $d_i = x_i - \overline{x}$, então $\sum_{i=1}^{n} d_i = 0$.

pp. 188 e 189

- $\underset{\sim}{x} = (2, 4, 6, 8)$; $\overline{x} = \dfrac{20}{4} = 5$
- $\underset{\sim}{y} = (3, 5, 7, 9)$; $\overline{y} = 5 + 1 = 6$
- $\underset{\sim}{z} = (6, 12, 18, 24)$; $\overline{z} = 3 \times 5 = 15$
- $\underset{\sim}{w} = (7, 13, 19, 24)$; $\overline{w} = 3 \times 5 + 1 = 16$

Soma dos quadrados dos desvios

- $SS_x = \sum_{i=1}^{n} (x_i - \overline{x})^2$
- $SS_x = \sum_{i=1}^{n} {x_i}^2 - n\overline{x}^2$

p. 192

$\underset{\sim}{x} = (2, 0, 4)$, tal que $\overline{x} = 2$.

$SS_x = (2-2)^2 + (0-2)^2 + (4-2)^2 =$

$= 0^2 + 4 + 4 = 8$

Propriedades de SS_x

- $SS_x = 0$ se e somente se $x_1 = x_2 = \ldots = x_n$.
- Para $h \neq 0$ e $n \in \mathbb{N}$:
  $\underset{\sim}{x} = (x_1, x_2, \ldots, x_n)$
  $\underset{\sim}{y} = (x_1 + h, x_2 + h, \ldots, x_n + h)$, então $SS_y = SS_x$.
- Para $a > 0$ e $n \in \mathbb{N}$:
  $\underset{\sim}{x} = (x_1, x_2, \ldots, x_n)$
  $\underset{\sim}{y} = (ax_1, ax_2, \ldots, ax_n)$, então $SS_y = a^2 SS_x$.

pp. 194 e 195

Sejam:

$\underset{\sim}{x} = (2, 3, 6, 1)$; $\quad \overline{x} = 3$

$\underset{\sim}{y} = (3, 4, 7, 2)$; $\quad \overline{y} = 4$

$\underset{\sim}{z} = (10, 15\ 30, 5)$; $\quad \overline{z} = 15$

- $SS_y = SS_x$
- $SS_z = 25 SS_x$

Variância e desvio-padrão de uma amostra

Seja a amostra $\underset{\sim}{x} = (x_1, x_2, \ldots x_n)$:

- **Variância**: $s_x{}^2 = \dfrac{SS_x}{n-1}$
- **Desvio-padrão**: $s_x = \sqrt{\dfrac{SS_x}{n-1}}$

p. 197

Seja $\underset{\sim}{x} = (2, 3, 4)$, tal que $\overline{x} = 2$.

- $SS_x = 8$
- $s_x{}^2 = \dfrac{8}{2} = 4$
- $s_x = \sqrt{4} = 2$

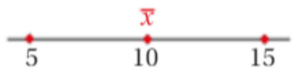
Média e desvio-padrão como medidas de dispersão

Dada uma amostra $(x_1, x_2, \ldots, x_n)$ de desvio-padrão não nulo, para qualquer k positivo, a percentagem de unidades estatísticas com valor fora do intervalo $[\overline{x} - ks_x, \overline{x} + ks_x]$ é sempre menor que $\dfrac{1}{k^2}$.

p. 200

Seja A uma amostra tal que:

$\overline{x} = 10$, $s_x = 2{,}5$ e $k = 2$



$\alpha < \dfrac{1}{k^2} = \dfrac{1}{4} = 25\%$

Não há mais de 25% das unidades estatísticas de A com valor fora do intervalo $[5, 15]$.

Percentis

O percentil de ordem k da amostra $\underset{\sim}{x} = (x_1, x_2, \ldots, x_n)$ é:

- o valor máximo da amostra se $k = 100$;
- a média dos elementos de ordem $\dfrac{kn}{100}$ e $\dfrac{kn}{100} + 1$ na amostra ordenada se $k \neq 100$ e $\dfrac{kn}{100}$ for inteiro;
- o elemento de ordem $\left[\dfrac{kn}{100}\right] + 1$ na amostra nos restantes casos.

Para dados agrupados, o percentil P_k determina-se utilizando o histograma correspondente.

pp. 201 a 203

Seja A a amostra ordenada:

$$(110, 150, 180, 200)$$

- $\boldsymbol{P_{30} = ?}$

  $\dfrac{5 \times 30}{100} = 1{,}5$ (não é um número inteiro)

  $P_{30} = [1{,}5] + 1 = 2$

  $P_{30} = x_{(2)} = 150$

- $\boldsymbol{P_{80} = ?}$

  $\dfrac{5 \times 80}{100} = 4$ (número inteiro)

  $P_{80} = \dfrac{x_{(3)} + x_{(4)}}{2} = \dfrac{180 + 200}{2} = 190$