Medidas de Dispersão

Medidas de Dispersão

O resumo de um conjunto de dados por uma única medida representativa de posição esconde informações sobre a variabilidade do conjunto de observações

Tipos

As principais medidas de dispersão absolutas são: amplitude total, desvio médio simples, variância e desvio padrão.

Medidas de Dispersão

- Por exemplo:

Grupos de alunos	notas	média
A	3, 4, 5, 6, 7	5
B	1, 3, 5, 7, 9	5
C	5, 5, 5, 5, 5	5
D	3, 5, 5, 7	5
E	2, 4, 5, 5, 9	5

Amplitude

É a diferença entre o maior e o menor valor da seqüência

Cálculo da amplitude

- Variável discreta
  - A amplitude total é a diferença entre o primeiro e o último elemento da série.
  - Exemplo:

Grupos de alunos	Notas	Média	Amplitude
A	3, 4, 5, 6, 7	5	4
B	1, 3, 5, 7, 9	5	8
C	5, 5, 5, 5, 5	5	0
D	3, 5, 5, 7	5	4
E	2, 4, 5, 5, 9	5	7

Cálculo da amplitude

- Variável contínua
  - A amplitude total é a diferença entre o ponto médio da última classe e o ponto médio da primeira classe.
  - Exemplo:

Classes	sálarios	Ponto médio	Freqüêcia
1	2 ├─ 4	3	5
2	4 ├─ 6	5	10
3	6 ├─ 8	7	20
4	8 ├─ 10	9	7
5	10 ├─ 12	11	2

Amplitude = 11 – 3 = 8

Desvio médio simples

O desvio médio simples que indicamos por DMS é definido como sendo uma média aritmética dos desvios de cada elemento da série para a média da série.

Cálculo dos desvio médio simples

- Variável discreta
  - A fórmula para o cálculo do DMS é

$$DMS = \frac{\sum_{i=1}^{k} n_i \,|\, x_i - \bar{x} \,|}{n} = \sum_{i=1}^{k} f_i \,|\, x_i - \bar{x} \,|$$

Onde $\bar{x}$ é a média da série, x_i é o valor da variável, n_i é a frequência, f_i é a frequência relativa e n somatório das frequências.

Cálculo dos desvio médio simples

- Variável contínua
  - Neste caso, por desconhecer os valores individuais dos elementos da série, substituindo os valores x_i pelos pontos médios de cada classe xp_i .

$$DMS = \frac{\sum_{i=1}^{k} n_i \, | \, xp_i - \bar{x} \, |}{n} = \sum_{i=1}^{k} f_i \, | \, xp_i - \bar{x} \, |$$

Variância e desvio padrão amostral

- A variância e o desvio padrão medem a dispersão dos dados em torno de sua média.
  - Variância: $$\mathrm{var}(X) = s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$
  - Desvio padrão: $$dp(X) = s = \sqrt{\mathrm{var}(X)}$$

Variância e desvio padrão populacional

- A variância e o desvio padrão medem a dispersão dos dados em torno de sua média.
  - Variância: $$\mathrm{var}(X) = \sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}$$
  - Desvio padrão: $$dp(X) = \sigma = \sqrt{\mathrm{var}(X)}$$

Coeficiente de Variação

- O coeficiente de variação é igual ao desvio-padrão dividido pela média aritmética, multiplicado por 100%.
- CV = $\left(\frac{s}{\bar{x}}\right) 100\%$
- É bastante útil para comparar dois ou mais conjuntos de dados que são mensurados em unidades diferentes.

Escores Z

- O escore Z de um dado corresponde à diferença entre o valor e a média, dividida pelo desvio-padrão.
- $Z = \frac{x - \bar{x}}{s}$

Exemplo: considerando amostras

Grupos de alunos	notas	média
A	3, 4, 5, 6, 7	5
B	1, 3, 5, 7, 9	5
C	5, 5, 5, 5, 5	5
D	3, 5, 5, 7	5
E	3, 5, 5, 6, 5	5

variância	Desvio padrão
2,5	1,58
10	3,16
0	0
2,667	1,63
1,667	1,29

Exercício 1

Uma pesquisa sobre consumo de gasolina deu os seguintes valores para a quilometragem percorrida por três marcas de carro (de mesma classe), em cinco testes com um tanque de 40 l:

Carro A	400	397	401	389	403
Carro B	403	401	390	378	395
Carro C	399	389	403	387	401

Compare as três marcas, qual a mais confiável.

Exercício 2

O que acontece com a mediana, a média e o desvio padrão de uma série de dados quando:

a) Cada observação é multiplicada por 2;

b) Soma-se 10 a cada observação;

c) Subtrai-se a média geral x de cada observação;

d) De cada observação subtrai-se x e divide-se pelo desvio padrão.