

Estatística Descritiva

Julienne Borges

Medidas de dispersão

Pra quê?

10 20 30 40 50

Média= 30

Mediana= 30

30 30 30 30 30

Média= 30

Mediana= 30

As amostras são iguais
em termos de medidas
de tendência central
mas, e em relação à
VARIABILIDADE??

Amplitude

- A notação utilizada para representar a amplitude é: A.
- É definida como sendo a diferença entre o maior e o menor valor do conjunto de dados.
- A vantagem da amplitude é sua facilidade de cálculo porém, tem a desvantagem de levar em conta apenas dois valores, desprezando todos os outros.

Variância e Desvio Padrão

- A *variância* e o *desvio padrão* medem a variação do conjunto de dados em torno da média.
- As notações utilizadas para representarmos as duas medidas são:
 - σ^2 : Variância populacional (Parâmetro)
 - s^2 : Variância amostral (Estatística)

Variância e Desvio Padrão

σ : Desvio padrão populacional (Parâmetro)

s : Desvio padrão amostral (Estatística)

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

Cálculo da variância
amostral

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$$

Cálculo da variância
populacional

Variância e Desvio Padrão

- É possível perceber que a unidade de medida da variância equivale à unidade de medida dos dados ao quadrado. Dessa maneira, é mais comum trabalharmos com a raiz quadrada da variância, ou seja, com o *desvio padrão*.

Exemplo 1

Considere uma amostra do tempo (em minutos) que funcionários de uma empresa gastam para chegar ao trabalho:

5 12 25 43 67

Para calcular o desvio padrão de tais medidas precisamos primeiramente calcular a média amostral:

Exemplo 1

$$\bar{x} = \frac{5 + 12 + 25 + 43 + 67}{5} = 30,4$$

Seguimos com o cálculo do desvio padrão (raiz quadrada da variância):

$$s = \sqrt{\frac{(5 - 30,4)^2 + (12 - 30,4)^2 + (25 - 30,4)^2 + (43 - 30,4)^2 + (67 - 30,4)^2}{5 - 1}}$$

$$s = 25,0559 \text{ minutos}$$

Exemplo 1

$$\bar{x} = \frac{5 + 12 + 25 + 43 + 67}{5} = 30,4$$

Seguimos com o cálculo do desvio padrão (raiz quadrada da variância):

$$s = \sqrt{\frac{(5 - 30,4)^2 + (12 - 30,4)^2 + (25 - 30,4)^2 + (43 - 30,4)^2 + (67 - 30,4)^2}{5 - 1}}$$

$$s = 25,0559 \text{ minutos}$$

Exemplo 2

Considere as notas de três turmas em uma prova de 10 pontos:

Turma	n	Média	Desvio padrão
A	10	5	1,5
B	10	5	3
C	10	5	4,5

Qual turma apresentou maior variabilidade nas notas?

Turma	n	Média	Desvio padrão
D	10	3	1,5
E	10	8	3
F	10	5	4,5

Qual turma apresentou maior variabilidade nas notas?

Coeficiente de Variação

- O desvio padrão embora seja a medida de dispersão mais utilizada, ela mede a dispersão em termos absolutos.
- O coeficiente de variação definido por

$$CV = \frac{s}{\bar{x}} \cdot 100$$

mede a dispersão em termos relativos.

Coeficiente de Variação

- O coeficiente de variação torna-se útil quando queremos comparar a variabilidade de observações com diferentes unidades de medidas.

Exemplo 2

Voltando aos dados do Exemplo 2:

Turma	Média	Desvio padrão	CV
A	5	1,5	30%
B	5	3	60%
C	5	4,5	90%

Qual turma apresentou maior variabilidade nas notas? E maior homogeneidade?

Turma	Média	Desvio padrão	CV
D	3	1,5	50%
E	8	3	37,5%
F	5	4,5	90%

Qual turma apresentou maior variabilidade nas notas? E maior homogeneidade?



PUC Minas
Virtual