By @kakashi_copiador

RESUMO DA AULA

MEDIDAS DE VARIABILIDADE

As medidas de dispersão (ou variabilidade) são justamente métricas que mostram a variação dos dados de um conjunto. Elas podem ser divididas em dois grupos:

Medidas de dispersão ABSOLUTA. → Amplitude Total; Amplitude Interquartílica; Desvio Médio; Variância; Desvio Padrão.

Medidas de dispersão RELATIVA. → Coeficiente de Variação de (Pearson); Variância Relativa.

AMPLITUDE TOTAL

A **amplitude total** é a diferença entre o **maior** e o **menor** elemento de um conjunto:

$$A = x_{máx} - x_{mín}$$

Sobre a amplitude total, podemos afirmar que:

I – No cálculo da amplitude total **desconsideramos os valores** da série que se encontram **entre os extremos**;

II – É **sensível ao tamanho de amostra**;

III – Pode apresentar muita **variação de uma amostra para outra,** ainda que extraídas de uma **mesma população.**

Amplitude Total para dados não-agrupados

Para dados não agrupados, o cálculo da amplitude total pode ser expresso pela seguinte fórmula:

$$\boxed{A = x_{máx} - x_{mín}}$$

Amplitude Total para dados agrupados sem intervalos de classes

Para dados agrupados sem intervalos de classe, a fórmula usada para a identificação da amplitude total é similar à adotada para dados não-agrupados. A única diferença consiste na identificação dos valores mínimo e máximo, que agora ocorre por meio de uma tabela de frequências.

Amplitude Total para dados agrupados em classes

Para dados agrupados em intervalos de classe, podemos definir a amplitude total de duas formas:

Para dados agrupados em intervalos de classe, podemos definir a amplitude total de duas formas:

- pela diferença entre o limite superior da última classe (L_{sup}) e o limite inferior da primeira classe (l_{inf}):

$$\boxed{A = L_{sup} - l_{inf}}$$

- pela diferença entre o ponto médio da última classe ($PM_{últ}$) e o ponto médio da primeira classe (PM_{pri}):

$$\boxed{A = PM_{últ} - PM_{pri}}$$

Propriedades da Amplitude Total

1ª Propriedade

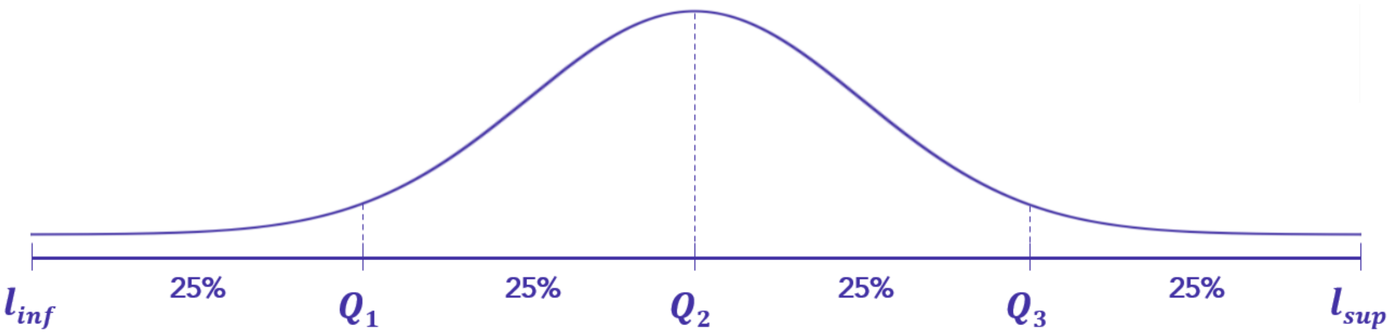
- **Somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, a amplitude do conjunto não é alterada.**

2ª Propriedade

- **Multiplicando-se (ou dividindo-se) todos os valores de uma variável por uma constante c, a amplitude do conjunto fica multiplicada (ou dividida) por essa constante.**

AMPLITUDE INTERQUARTÍLICA

Denominamos de quartis os valores de uma série que a dividem em quatro partes iguais, isto é, quatro partes contendo o mesmo número de elementos (25%).



A amplitude interquartílica é o resultado da subtração entre o terceiro quartil e o primeiro quartil:

$$\boxed{A_{IQ} = Q_3 - Q_1}$$

A amplitude semi-interquartílica é definida como a metade desse valor, sendo calculada pela expressão:

$$\boxed{D_Q = \frac{Q_3 - Q_1}{2}}$$

Propriedades da Amplitude Interquartílica

1ª Propriedade

- **Somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, a amplitude interquartílica (e o desvio quartílico) do conjunto não é alterada.**

2ª Propriedade

- **Multiplicando-se (ou dividindo-se) todos os valores de uma variável por uma constante c, a amplitude interquartílica (e o desvio quartílico) do conjunto fica multiplicada (ou dividida) por essa constante.**

DESVIOS EM RELAÇÃO À MÉDIA ARITMÉTICA E MEDIANA

Um desvio é a distância entre qualquer observação do conjunto de dados e uma medida descritiva desse conjunto:

$$desvio = observação - medida$$

Em especial, destacamos os desvios em relação à média aritmética e em relação à mediana:

$$\boxed{\boldsymbol{d_i = x - \overline{x}} \quad (média)}$$

ou

$$\boxed{\boldsymbol{d_i = x - M_d} \quad (mediana)}$$

Propriedades dos Desvios em Relação à Média Aritmética e Mediana

1ª Propriedade

- **A soma dos desvios em relação à média aritmética é sempre nula.**

2ª Propriedade

- **A soma dos quadrados dos desvios em relação à média aritmética é mínima.**

3ª Propriedade

- **A soma dos módulos dos desvios em relação à mediana é mínima.**

DESVIO ABSOLUTO MÉDIO

DESVIO ABSOLUTO MÉDIO

mede a dispersão entre os valores da distribuição e a média dos dados coletados.

$$D_m = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

Desvio Médio para dados não-agrupados

$$D_m = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{n}$$

Desvio Médio para dados agrupados sem intervalo de classe

$$D_m = \frac{\sum_{i=1}^{m} |x_i - \bar{x}| \times f_i}{\sum f_i}$$

Desvio Médio para dados agrupados em classes

Para dados agrupados em classe, deveremos adotar a mesma convenção que tomamos para o cálculo da média: vamos assumir que todos os valores coincidem com os pontos médios das suas respectivas classes.

Propriedades do Desvio Médio

1ª Propriedade

- **Somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, o desvio médio do conjunto não é alterado.**

2ª Propriedade

- **Multiplicando-se (ou dividindo-se) todos os valores de uma variável por uma constante c, o desvio médio do conjunto fica multiplicado (ou dividido) por essa constante.**

VARIÂNCIA (σ^2)

Fórmula da variância populacional:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n} \quad \text{ou} \quad \sigma^2 = \overline{x^2} - (\bar{x})^2$$

Fórmula da variância amostral:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1} \quad \text{ou} \quad s^2 = \left[\overline{x^2} - (\bar{x})^2\right] \times \left(\frac{n}{n-1}\right)$$

Variância para dados não-agrupados

a) para **populações**:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n} \quad ou \quad \sigma^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n}$$

b) para **amostras**:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n - 1} \quad ou \quad s^2 = \frac{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}}{n - 1}$$

A relação entre a **variância amostral (s^2) e a variância populacional (σ^2)** é dada por:

$$s^2 = \left(\frac{n}{n - 1}\right) \times \sigma^2$$

Variância para dados agrupados sem intervalos de classes

a) para **populações**:

$$\sigma^2 = \frac{\sum_{i=1}^{m}(X_i - \mu)^2 \times f_i}{n} \quad ou \quad \sigma^2 = \frac{\sum_{i=1}^{m}\left(X_i^2 \times f_i\right) - \frac{\left(\sum_{i=1}^{m} X_i \times f_i\right)^2}{n}}{n}$$

b) para **amostras**:

$$s^2 = \frac{\sum_{i=1}^{m}(X_i - \overline{x})^2 \times f_i}{n - 1} \quad ou \quad s^2 = \frac{\sum_{i=1}^{m}\left(X_i^2 \times f_i\right) - \frac{\left(\sum_{i=1}^{m} X_i \times f_i\right)^2}{n}}{n - 1}$$

Variância para dados agrupados em classes

a) para populações:

$$\sigma^2 = \frac{\sum_{i=1}^{k}(PM_i - \mu)^2 \times f_i}{n} \quad ou \quad \sigma^2 = \frac{\sum_{i=1}^{k} PM_i^2 \times f_i - \frac{\left(\sum_{i=1}^{k} PM_i \times f_i\right)^2}{n}}{n}$$

b) para amostras:

$$s^2 = \frac{\sum_{i=1}^{k}(PM_i - \overline{x})^2 \times f_i}{n - 1} \quad ou \quad s^2 = \frac{\sum_{i=1}^{k} PM_i^2 \times f_i - \frac{\left(\sum_{i=1}^{k} PM_i \times f_i\right)^2}{n}}{n - 1}$$

Propriedades da Variância

1ª Propriedade

- **Somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, a variância do conjunto não é alterada.**

2ª Propriedade

- **Multiplicando-se (ou dividindo-se) todos os valores de uma variável por uma constante c, a variância do conjunto fica multiplicada (ou dividida) pelo QUADRADO dessa constante.**

DESVIO-PADRÃO (σ)

Fórmula do desvio-padrão populacional:

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

Fórmula do desvio-padrão amostral:

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

Desvio-padrão para dados não-agrupados

a) para populações:

$$\boldsymbol{\sigma = \sqrt{\frac{\sum_{i=1}^{n} d_i^2}{n}} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \mu)^2}{n}}}$$

b) para amostras:

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

Desvio-padrão para dados agrupados sem intervalo de Classe

a) para populações:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{k}\left(d_i^2 \times f_i\right)}{n}} = \sqrt{\frac{\sum_{i=1}^{k}[(X_i - \mu)^2 \times f_i]}{n}}$$

b) para amostras:

$$s = \sqrt{\frac{\sum_{i=1}^{k}\left(d_i^2 \times f_i\right)}{n-1}} = \sqrt{\frac{\sum_{i=1}^{k}[(X_i - \overline{x})^2 \times f_i]}{n-1}}$$

Desvio-padrão para dados agrupados em classes

Para dados agrupados em classe, usaremos as mesmas fórmulas para dados sem intervalos de classes, utilizando para x_i os pontos médios de cada classe, mas adotando os mesmos procedimentos.

Propriedades do Desvio-padrão

1ª Propriedade

- **Somando-se (ou subtraindo-se) uma constante c a todos os valores de uma variável, o desvio-padrão do conjunto não é alterado.**

2ª Propriedade

- **Multiplicando-se (ou dividindo-se) todos os valores de uma variável por uma constante c, o desvio-padrão do conjunto fica multiplicado (ou dividido) por essa constante.**

COEFICIENTE DE VARIAÇÃO (OU DISPERSÃO RELATIVA)

O coeficiente de variação é uma medida que fornece a variação dos dados em relação à média:

a) para populações:

$$\boldsymbol{CV = \frac{\sigma}{\mu} \times 100\ (\%)}$$

b) para amostras:

$$\boldsymbol{CV = \frac{s}{\bar{x}} \times 100\ (\%)}$$

O coeficiente de variação pode ser interpretado por meio de algumas regras empíricas:

A DISTRIBUIÇÃO TEM BAIXA DISPERSÃO	A DISTRIBUIÇÃO É TEM MÉDIA DISPERSÃO	A DISTRIBUIÇÃO TEM ELEVADA DISPERSÃO
$CV < 15\%$	$15\% < CV < 30\%$	$CV > 30\%$

podemos classificar as distribuições em homogêneas ou heterogêneas:

A DISTRIBUIÇÃO HOMOGÊNEA QUANDO POSSUI DISPERSÃO BAIXA OU MÉDIA	A HETEROGÊNEA QUANDO POSSUI DISPERSÃO ELEVADA
$CV < 30\%$	$CV > 30\%$

VARIÂNCIA RELATIVA

A **variância relativa** é uma medida de dispersão relativa que resulta do quociente entre a **variância absoluta** e o **quadrado da média**:

a) para **populações**:

$$\boxed{VR = \left(\frac{\sigma}{\mu}\right)^2 = \frac{\sigma^2}{\mu^2}}$$

b) para **amostras**:

$$\boxed{VR = \left(\frac{s}{\bar{x}}\right)^2 = \frac{s^2}{\bar{x}^2}}$$