

Medidas de Dispersão

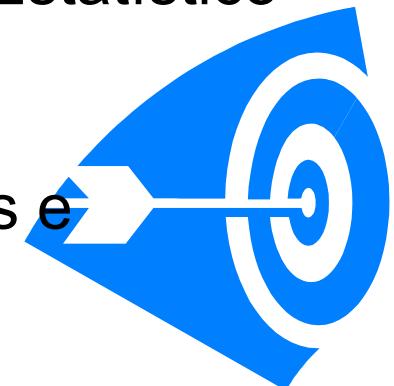
Introdução
Amplitude
Variância
Desvio Padrão
Coeficiente de Variação

Introdução

- Estudo de medidas que mostram a dispersão dos dados em torno da tendência central
- Analisaremos as seguintes medidas:
 - Amplitude
 - Variância
 - Desvio Padrão
 - Coeficiente de Variação

Introdução

- Por que estudar?
 - Entender grandes volumes de informação (pesquisas de mercado, índices populacionais, acessos a sites)
 - Qualidade de processos (CEP – Controle Estatístico de Processos)
 - Previsões confiáveis (projeções financeiras e populacionais, vida útil de equipamentos)
 - Planejamento (coleta de dados, definição de amostras, planos de contingência)



Amplitude

- É a diferença entre a maior e a menor observação em um conjunto de dados

$$A = X_{maior} - X_{menor}$$

- Mede a dispersão total no conjunto de dados
- É uma medida simples que não leva em consideração como os dados são efetivamente distribuídos entre os valores extremos

Amplitude

- Exemplo: Dada a amostra abaixo, calcule a amplitude.

Custo de produção (em milhões)
2
3
3,5
3,7
5
12

$$A = 12 - 2 = 10$$

OBS: A amplitude calculada não nos fornece qualquer informação sobre a tendência central e distribuição das observações

Variância

- A variância da amostra é a média aproximada das diferenças ao quadrado entre cada uma das observações e a média aritmética da amostra

$$S^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + (X_3 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n - 1}$$

onde:

n é o tamanho da amostra

OBS: O tamanho da amostra é subtraído de 1 devido ao fator de correção de Bessel, que visa uma estimativa mais precisa. No cálculo de variância para toda a população, este corretor é dispensado.

Variância

- A fórmula da variância de uma amostra pode ser escrita de forma resumida

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}$$

Resumindo: A variância é a soma das diferenças ao quadrado em torno da média aritmética dividida pelo tamanho da amostra menos um

Variância

- A variância também pode ser calculada pela fórmula a seguir, que exige um número menor de operações aritméticas

$$S^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}$$

- A variância da população é representada pelo símbolo σ^2 , porém é mais comum e prático o cálculo da variância da amostra

Variância

- Exercício: Calcule a variância da amostra

Idades

16
42
20
22
35
18

Variância

- Exemplo: Calcule a variância da amostra

x_i	$x_i - \bar{x}$	d_i^2
16	-9,5	90,25
42	16,5	272,25
20	-5,5	30,25
22	-3,5	12,25
35	9,5	90,25
18	-7,5	56,25

$$\bar{X} = 25,5$$

$$\sum d_i^2 = 551,5$$

$$S^2 = \frac{551,5}{5} \approx 110,3$$

Variância

- Propriedades
 1. Somando-se (ou subtraindo-se) a cada elemento de um conjunto de valores uma constante arbitrária, a variância não se altera
 2. Multiplicando-se (ou dividindo-se) cada elemento de um conjunto de valores por um valor constante, a variância fica multiplicada (ou dividida) pelo quadrado da constante

Desvio Padrão

- Desvio padrão é a raiz quadrada da variância da amostra

$$S = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n - 1}}$$

Desvio Padrão

- O desvio padrão indica o afastamento dos valores observados em relação à média aritmética da amostra estudada
- É um conceito imprescindível para análises gráficas, determinação de confiabilidade e estudos de distribuições

Desvio Padrão

- Exemplo: com base na amostra utilizada no exercício de variância, calcule o desvio padrão.

$$S = \sqrt{S^2} = \sqrt{\frac{551,5}{5}} = 10,5$$

Desvio Padrão

- Propriedades
 1. Somando-se (ou subtraindo-se) a cada elemento de um conjunto de valores uma constante arbitrária, o desvio padrão não se altera
 3. Multiplicando-se (ou dividindo-se) cada elemento de um conjunto de valores por um valor constante, desvio padrão fica multiplicado (ou dividido) pela constante

Coeficiente de Variação

- Dentre as medidas de dispersão relativa, o coeficiente de variação mais usado é o de Pearson (CV_p)
- Pode ser expresso em porcentagem (taxa de variação)
- Útil para comparação de variabilidade de dois conjuntos de dados com unidades de medidas diferentes
- Também é útil para comparar amostras de comportamento bastante diferentes (ex: ações de uma indústria X ações de empresa de serviços aéreos)

Coeficiente de Variação

- CV_p é baseado no quociente entre o desvio padrão e a média aritmética

$$CV_p = \left(\frac{S}{\bar{X}} \right) 100\%$$

- Quanto menor este valor, mais homogêneo será o conjunto de dados

Exercícios

- Para os dados usados nos exercícios de variância e desvio padrão, calcule o coeficiente de variação
- O desvio padrão e a variância podem ser negativos?
- Em que situação o desvio padrão e a variância são nulos? Qual é a amplitude neste caso?