

Medidas de Dispersão

ESTAT0011 – Estatística Aplicada

Prof. Dr. Sadraque E. F. Lucena

sadraquelucena@academico.ufs.br

Medidas de Dispersão

- Dois sistemas internos de uma empresa foram monitorados por 10 semanas, registrando o número de acessos nas sextas-feiras.

Semana	1	2	3	4	5	6	7	8	9	10
Servidor A	56	56	57	58	61	63	63	67	67	67
Servidor B	33	42	48	52	57	67	67	77	82	90

Nota

Média: 61,5 Mediana: 62 Moda: 67 (para ambos os servidores)

- A média, a mediana e a moda são iguais. Mas evidentemente, os dois sistemas não se comportam da mesma forma.

Medidas de Dispersão

- As medidas de posição apresentadas fornecem a informação dos dados apenas a nível pontual, sem ilustrar outros aspectos referentes à forma como os dados estão distribuídos na amostra.
- É preciso usarmos medidas estatísticas complementares para caracterizar melhor os dados apresentados.
- **Medidas de Dispersão (ou medidas de variação)** servem para caracterizar o quanto os dados estão espalhados em torno de uma medida de posição como, por exemplo, a média.

Na aula de hoje veremos...

- Amplitude total (AT);
- Variância (S^2);
- Desvio-padrão (S);
- Coeficiente de variação (CV).

Amplitude Total

- É a diferença entre o maior e o menor valor do conjunto de dados, ou seja,

$$AT = \text{máx} - \text{mín.}$$



Quando usar

- Comparações rápidas.
- Quando você quer saber o “campo de variação” do dados.



Limitações

- Sensível a extremos.
- Não considera a distribuição interna dos dados.

Variância

- Média dos quadrados dos desvios em relação à média.
- Mostra o **quanto os valores se afastam da média**.
- O cálculo muda para dados populacionais ou dados amostrais:

Dados	Variância
População	$\sigma^2 = \frac{\sum_{i=1}^N (X_i - \mu)^2}{N}$
Amostra	$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$

- Lembre que μ denota a média populacional e \bar{X} representa a média amostral.

Variância



Quando usar

- Para cálculos estatísticos mais avançados.
- Quando não importa que a unidade esteja ao quadrado.



Limitações

- Unidade dos dados é alterada (fica ao quadrado).
- Difícil interpretação direta.

Exemplo 3.1

Calcule a variância do número de acessos ao Servidor B e compare com a variância do servidor A ($\sigma_A^2 = 20,94 \text{ acessos}^2$).

Comentários Importantes

- Apesar de útil para descrever a variabilidade do conjunto de observações, o fato mencionado torna um pouco inviável a análise, dado que a unidade de medida fica elevada ao quadrado.
- O interessante é ter uma medida que descreva a variabilidade das informações na mesma escala em que estão os dados fornecidos.
- Esta medida se chama **desvio padrão**.

Desvio Padrão

- **Definição:** raiz quadrada da variância.

$$S = \sqrt{S^2}$$

- Prefere-se usar o desvio padrão porque este é expresso na mesma unidade dos dados.



Quando usar

- Quando quiser uma medida de dispersão **na mesma unidade dos dados**.
- Mais interpretável que a variância.

Exemplo 3.2

Qual o desvio padrão o número de aceso dos dois servidores do Exemplo 3.1?

Coeficiente de Variação

- É uma medida relativa de dispersão.
- Mostra a dispersão proporcional dos dados em relação à média:

$$CV = \frac{S}{\bar{X}} \quad \text{ou} \quad CV\% = \frac{S}{\bar{X}} \times 100$$



Quando usar

- Para **comparar dispersões entre conjuntos com médias diferentes ou unidades diferentes**.
- Útil para **decisões sobre escalabilidade, risco ou eficiência*.

Exemplo 3.3

Uma equipe de desenvolvimento monitorou o tempo de resposta (em milissegundos) de duas APIs utilizadas internamente. Calcule o coeficiente de variação (CV) para cada API e indique qual API apresenta desempenho mais estável (o objetivo é evitar picos de latência que podem prejudicar a experiência do usuário).

Chamada	API A	API B
1	120	100
2	122	160
3	121	80
4	119	150
5	118	90

Aplicações práticas

Situação Prática	Medida Recomendada	Por quê?
Análise de acessos semanais a servidores	Desvio padrão / CV	Avaliar estabilidade do sistema / escalabilidade
Comparar desempenho de dois algoritmos	CV	Permite comparação mesmo com métricas diferentes
Identificar picos de uso	Amplitude	Ver o intervalo extremo de demanda
Avaliar consistência de tempos de resposta	Desvio padrão	Ver se o sistema responde com regularidade

Fim