


The background is a dark, blue-tinted image of an industrial factory floor. In the foreground, there are various mechanical parts and structures. In the background, there are large circular gauges or displays showing percentages: 100%, 80%, 79%, and 94%. The text 'UnISENAI' is prominently displayed in the center, with 'Uni' in blue and 'SENAI' in white with a blue glow effect.


UnISENAI

O FUTURO COMEÇA
POR VOCÊ!

The background is a dark blue-tinted image of a modern industrial factory floor. In the foreground, there's a large, complex machine, possibly a CNC lathe or mill. Overlaid on the image are several semi-transparent data visualizations: a line graph with a rising trend, a pie chart, and several circular gauges or progress indicators. One gauge shows '80%', another '79%', and a larger one in the center-right shows '94%'. The overall aesthetic is high-tech and data-driven.

Pós-graduação em Ciência de Dados e Inteligência Artificial

UniSENAI



Estatística Aplicada à Análises de Dados

Tópico 01 – Estatística Descritiva

UniSENAI

Estatística Descritiva – Continuação

1.5 MEDIDAS DESCRITIVAS OU MEDIDAS RESUMO

1.5 Medidas Descritivas ou Medidas Resumo

Medidas de Tendência Central



Medidas de tendência central são utilizadas na estatística para resumir ou descrever um conjunto de dados, indicando um valor que representa uma "tendência" no centro dos dados. Elas fornecem informações sobre a localização ou o ponto central dos dados. As medidas de tendência central são:

- Média
- Mediana
- Moda

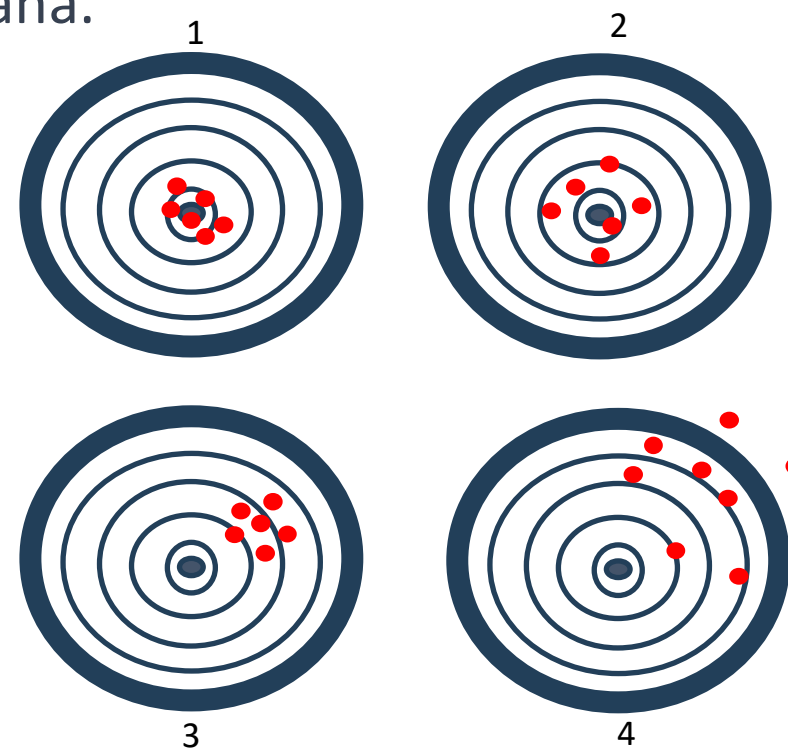
1.5 Medidas Descritivas ou Medidas Resumo

Medidas de Variabilidade

Medidas de variabilidade, também conhecidas como medidas de dispersão, são utilizadas na estatística para avaliar o grau de dispersão ou variabilidade dos dados em relação à medida de tendência central. Elas fornecem informações sobre o quão distantes ou espalhados os valores estão em relação à média ou à mediana.

As medidas de Dispersão são:

- Amplitude Total
- Desvio-Padrão
- Variância
- Coeficientes de Variação
- Desvio Interquartílico



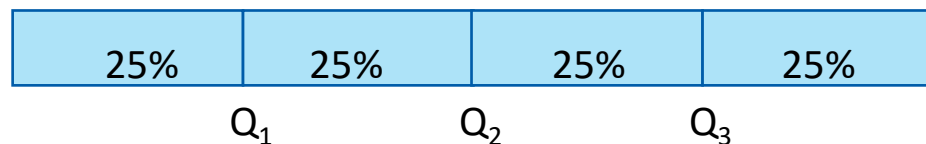
1.5 Medidas Descritivas

Medidas Separatrizes

Medidas separatrizes, também conhecidas como medidas de posição, são utilizadas na estatística para dividir ou separar um conjunto de dados em partes iguais ou em proporções específicas. Elas permitem identificar a posição relativa de um valor dentro de um conjunto de dados e são úteis para entender a distribuição dos dados e identificar valores atípicos.

As medidas de Posição são:

- Quartis – Dividem os dados em 4 partes (Quartil1, Quartil2 e Quartil3)



- Percentis – Dividem os dados em 100 partes iguais. Exemplo: Percentil 25 = Quartil1, Percentil 50 = Quartil2 = Mediana e o percentil 10 (elemento que separa os

Média

- Uma das medidas mais utilizadas para representar uma distribuição de dados
- Reflete um ponto de equilíbrio
- Influenciada por valores extremos
- Só deve ser utilizada em distribuições:
 - Não heterogêneas
 - Simétricas ou levemente assimétricas

Exemplo: 3,2 3,2 2,8 2,1 2,9 3,1 3,2 3,0 3,5 4,0

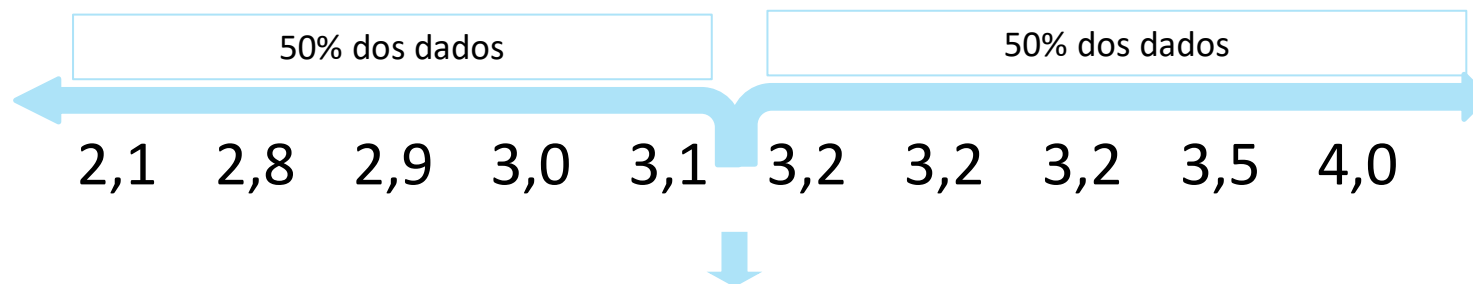
$$\bar{x} = \frac{3,2 + 3,2 + 2,8 + \dots + 3,5 + 4,0}{10} = \frac{\sum_{i=1}^n x_i}{n} = 3,1$$

Mediana

- A mediana é o valor que divide a distribuição em duas partes iguais
 - Se n for ímpar, a mediana será a obs. que ocupa a posição central
 - Se n for par, a mediana será a média dos dois valores centrais
- A mediana não é influenciada por valores extremos.

Exemplo: 1º Passo: Ordenar os dados

2º Passo: Dividir os dados ao meio



Mediana

$$\tilde{x} = \frac{3,1 + 3,2}{2} = 3,15$$

Moda

- Moda é o valor mais frequente do DataSet
- Não é afetada por valores extremos
- Pode haver mais de uma moda
- Pode ser utilizada para dados qualitativos e quantitativos

Exemplo:

2,1 2,8 2,9 3,0 3,1 3,2 3,2 3,2 3,5 4,0

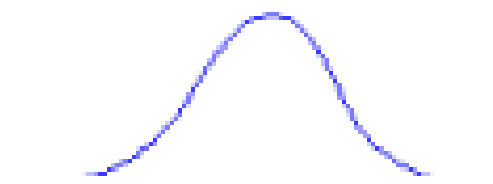


Moda = 3,2

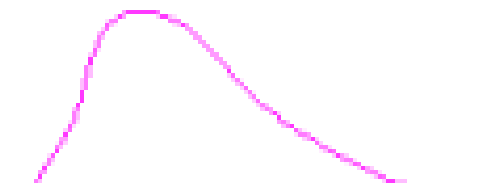
Quando temos uma tabela de frequências, o intervalo de classe com maior frequência será a classe modal.

A média é influenciada por outliers (valores muito grandes ou muito pequenos):

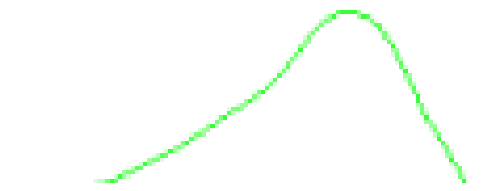
1. Distribuição simétrica \Rightarrow média = mediana
2. Distribuição assimétrica para a direita (alguns valores grandes como "outliers") \Rightarrow média > mediana
3. Distribuição assimétrica para a esquerda (alguns valores pequenos como "outliers") \Rightarrow média < mediana



média = mediana



média > mediana



média < mediana



a mediana do peso destes atletas é 75kg,
mas o peso médio é de 105kg !

Medidas de Variabilidade

Amplitude Total

- AT = Diferença entre a maior e a menor observação
- Ignora como os dados estão distribuídos
- Considera apenas dois valores da distribuição
- É a pior medida de variabilidade

Exemplo: 3,2 3,2 2,8 2,1 2,9 3,1 3,2 3,0 3,5 4,0

Amplitude Total

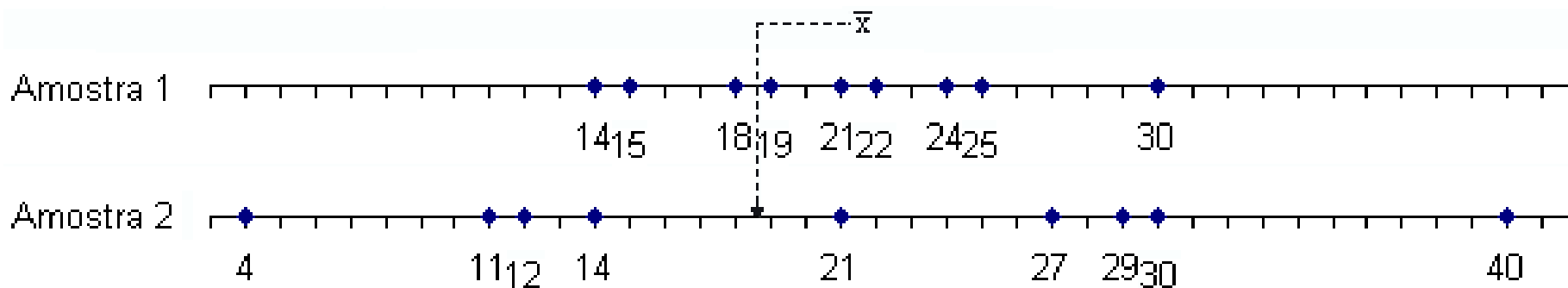
$$AT = \text{máximo} - \text{mínimo}$$

$$AT = 4 - 3,2 = 0,8$$

Medidas de Variabilidade

Variância e Desvio-Padrão

- Medidas mais comuns de variação absoluta
- Indicam como os dados estão concentrados em torno da média
- Quanto maior o valor do desvio-padrão (s) ou da variância, maior é a variabilidade dos dados
- No exemplo abaixo a amostra 2 apresenta desvio-padrão e variância maior que a amostra 1



$$\text{Variância} = S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \quad \Rightarrow \quad S^2 = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n - 1}$$

Desvio-Padrão (s)

$$S = \sqrt{s^2} \rightarrow \text{Desvio padrão} = \text{raiz da variância}$$

Exemplo:

Dados: 10 4 8 11 6 9

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1} \quad \text{onde} \quad \bar{x} = \frac{\sum_{i=1}^n x_i}{6} = 8$$

$$s^2 = \frac{(10-8)^2 + (4-8)^2 + \dots + (9-8)^2}{6-1} = 6,8$$

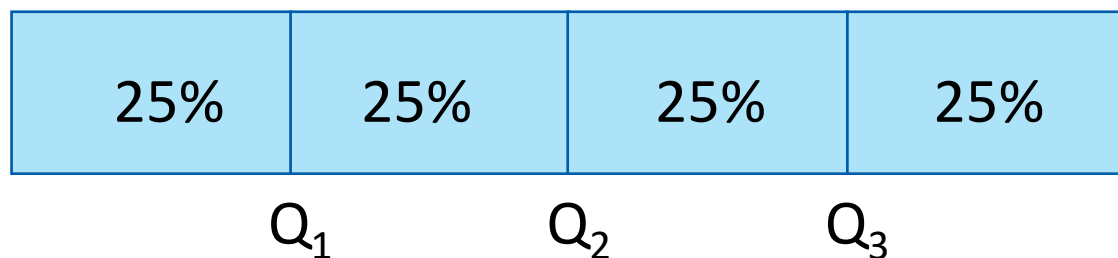
Desvio padrão = raiz de 6,8 = 2,6

Medidas Separatrizes ou de Posição

Quartis, Decis e Percentis


- 1º Quartil = Percentil 25 \Rightarrow 25% dos dados são $\leq Q_1$
- 2º Quartil = Percentil 50 \Rightarrow 50% dos dados são $\leq Q_2$ (mediana)
- 3º Quartil = Percentil 75 \Rightarrow 75% dos dados são $\leq Q_3$
- Os quartis separam os dados ordenados em 4 partes iguais
- Os Decis separam os dados ordenados em 10 partes iguais
- Os percentis separam os dados ordenados em 100 partes iguais
- Os quartis, além de descrever a amostra, são utilizados para a construção do Box-Plot

Exemplo: Se o percentil 75 da Variável peso de um DataSet é igual a 65,7 Kg, dizemos que 75% dos indivíduos do estudo apresentam peso menor ou igual a 65,7 Kg ou, que 25% tem pesos acima deste valor.



UniSENAI

Rodovia SC-401, 3730, Bairro Saco Grande, Florianópolis/SC

 3239 5745

unisenaisc.com.br



SENAI