

# Measures of Central Tendency and Variation

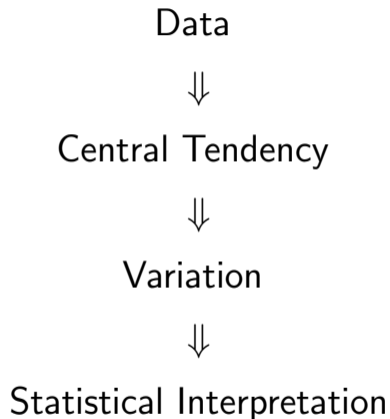
## Descriptive Statistics

Shahkar Ahmad Nahvi

DoEE, IUST

# Learning Outcomes

- explain measures of central tendency,
- compute mean, median and mode,
- understand measures of variation,
- compute range, variance and standard deviation,
- interpret statistical summaries correctly,
- apply concepts to real-life problems.



**A complete summary requires both center and spread.**

# 1. Measures of Central Tendency

Measures of central tendency describe the central or typical value of data.

Main measures:

- Mean
- Median
- Mode

Different measures are useful for different types of data.

## 2. Mean (Arithmetic Mean)

### Definition

$$\bar{x} = \frac{\sum x_i}{n}$$

### Example

Data: 5, 7, 9

$$\bar{x} = \frac{5 + 7 + 9}{3} = 7$$

### Applications

- Average marks
- Average temperature
- Business statistics

# Properties of Mean

- Uses all observations
- Easy to compute
- Widely used in statistics

## Limitation

Sensitive to extreme values (outliers).

## Example

Data: 5, 6, 7, 8, 100

$$\bar{x} = \frac{126}{5} = 25.2$$

Mean is distorted by the outlier 100.

## 3. Median

### Definition

Median is the middle value of ordered data.

### Procedure

- Arrange data in ascending order
- Odd  $n$ : middle value
- Even  $n$ : average of two middle values

### Example

Data: 2, 4, 6, 8

$$\text{Median} = \frac{4 + 6}{2} = 5$$

# Why Median is Important

Median is resistant to outliers.

## Example

Data: 5, 6, 7, 8, 100

Median = 7

Median better represents skewed data.

## Applications

- Income distribution
- Real estate pricing

## 4. Mode

### Definition

Mode is the most frequently occurring value.

### Example

Data: 2, 3, 3, 5

Mode = 3

### Applications

- Product demand
- Survey responses
- Shoe sizes

# Mean vs Median vs Mode

Feature	Mean	Median	Mode
Uses all data	Yes	No	No
Affected by outliers	Yes	No	No
Best for	Symmetric	Skewed	Frequency

No single measure is best for all situations

## 5. Measures of Variation

Measures of variation describe the spread or dispersion of data.

Main measures:

- Range
- Variance
- Standard Deviation
- Mean Absolute Deviation
- Coefficient of Variation

## 6. Range

### Definition

$$\text{Range} = \text{Maximum} - \text{Minimum}$$

### Example

Data: 2, 5, 9

$$\text{Range} = 9 - 2 = 7$$

### Limitation

Uses only two observations.

## 7. Variance

### Definition

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

### Example

Data: 2, 4, 6

Mean = 4

$$\sigma^2 = \frac{(2 - 4)^2 + (4 - 4)^2 + (6 - 4)^2}{3}$$

$$\sigma^2 = \frac{8}{3}$$

## 8. Standard Deviation

### Definition

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

### Example

$$\sigma = \sqrt{\frac{8}{3}} \approx 1.63$$

### Interpretation

Standard deviation measures average deviation from the mean.

Higher standard deviation means greater variability.

## 9. Mean Absolute Deviation (MAD)

### Definition

$$MAD = \frac{\sum |x_i - \bar{x}|}{n}$$

### Example

Data: 2, 4, 6

$$MAD = \frac{2 + 0 + 2}{3}$$

$$MAD = \frac{4}{3}$$

MAD uses absolute deviations instead of squared deviations.

## 10. Coefficient of Variation (CV)

### Definition

$$CV = \frac{\sigma}{\bar{x}} \times 100$$

### Example

$$CV = \frac{1.63}{4} \times 100 \approx 40.75\%$$

### Use

Compares variability across datasets with different scales.

# Variation Measures Comparison

Measure	Uses	Limitation
Range	Quick spread	Uses only extremes
Variance	Mathematical analysis	Units squared
SD	Practical interpretation	Sensitive to outliers
MAD	Simpler interpretation	Less widely used
CV	Relative variability	Requires nonzero mean

## 11. Solved Example: Outlier Effect

### Data

5, 6, 7, 8, 100

### Mean

$$\bar{x} = 25.2$$

### Median

$$\text{Median} = 7$$

Median better represents central tendency because of the outlier.

## 12. Solved Example: Missing Value

Mean of 5 numbers is 10.

Known numbers:

8, 9, 10, 11

### Step 1

Total sum:

$$5 \times 10 = 50$$

### Step 2

Known sum:

$$8 + 9 + 10 + 11 = 38$$

**Missing value**

# 13. Real-Life Applications

## **Economics**

- Median income

## **Education**

- Mean marks
- Standard deviation of scores

## **Business**

- Mode for demand analysis

## **Finance**

- Standard deviation for investment risk

## 14. Case Study: Income Distribution

Suppose a company has salaries:

*25k, 30k, 35k, 40k, 500k*

### **Issue**

Mean salary becomes misleading because of one extremely high salary.

### **Better Measure**

Median better represents typical income.

## 15. Case Study: Exam Difficulty

Two exams have same average score.

Exam A:

$$SD = 5$$

Exam B:

$$SD = 15$$

### **Interpretation**

Exam B has much greater spread in scores and may be less consistent.

## 16. Case Study: Product Demand

A shoe store records customer demand:

7, 8, 8, 8, 9, 10

**Mode**

8

**Business Insight**

Size 8 should have higher inventory.

## 17. Important Definitions

**Mean:** Arithmetic average.

**Median:** Middle value of ordered data.

**Mode:** Most frequent value.

**Variance:** Average squared deviation.

**Standard Deviation:** Square root of variance.

## 18. Numerical Example

Data:

10, 12, 14

**Mean**

$$\bar{x} = 12$$

**Variance**

$$\sigma^2 = \frac{8}{3}$$

**Standard Deviation**

$$\sigma \approx 1.63$$

# 19. Concept Check

## Question 1

Why is mean affected by outliers?

## Question 2

When is median preferred over mean?

## Question 3

What does a high standard deviation indicate?

## 20. Key Takeaways

- Mean uses all observations.
- Median is robust to outliers.
- Mode identifies popularity.
- Variation measures spread.
- Standard deviation measures uncertainty.
- Central tendency alone is insufficient.

# Thank You