جامعة المستقبل
AL MUSTAQBAL UNIVERSITY

Lecture 2

Descriptive Statistics

Mean, Median, Mode, Range, and Variance

By
Hawraa Aead AL qasear

Objectives

- By the end of this lecture, students will be able to:
- • Define descriptive statistics.
- • Identify measures of central tendency and dispersion.
- • Calculate and interpret mean, median, mode, range, and variance.
- • Understand how these measures summarize data.

Introduction

- Descriptive statistics help us summarize and describe data.
- They allow us to understand data patterns and trends.

- **Two main types:**
- • Measures of Central Tendency – show where the center of data lies.
- • Measures of Dispersion – show how spread out the data are.

Mean (Arithmetic Average)

- **Definition:**
- The mean is the sum of all data values divided by the number of values.

- Formula:
- **Mean = (Σx) / N**

- **Example:**
- Scores = 10, 20, 30, 40, 50
- Mean = (10 + 20 + 30 + 40 + 50) / 5 = 30

- Advantages:
- • Easy to calculate
- • Uses all data values

Median (Middle Value)

- Definition:
- The median is the middle value when data are arranged in ascending order.
- If there are two middle values, take their average.

- Examples:
- • 5, 10, 15 → Median = 10
- • 5, 10, 15, 20 → Median = (10 + 15)/2 = 12.5

- Advantages:
- • Not affected by outliers
- • Useful for skewed data

Mode (Most Frequent Value)

- Definition:
- The mode is the value that occurs most often.

- Example:
- Data = 2, 3, 3, 4, 5 → Mode = 3

- Notes:
- • If two values occur equally often → bimodal.
- • If all values are unique → no mode.

- Advantages:
- • Useful for categorical or nominal data (e.g., favorite color, brand).

Comparison: Mean, Median, and Mode

- • Mean – Best for data without extreme values.
- • Median – Best for skewed data or when outliers exist.
- • Mode – Best for qualitative or categorical data.

- Example:
- Data = 10, 10, 20, 100
- Mean = 35
- Median = 15
- Mode = 10

Range (Measure of Spread)

- Definition:
- Range = Maximum value - Minimum value

- Example:
- Scores = 10, 20, 30, 40, 50
- Range = 50 - 10 = 40

- Advantages:
- • Easy to calculate

- Disadvantages:
- • Considers only two values (max and min)
- • Does not show internal variation

Variance (Measure of Variability)

- Definition:
- Variance measures how far data values are from the mean.

- Formula:
- Variance (σ^2) = $\Sigma(x - \text{mean})^2 / N$

- Example:
- Data = 2, 4, 6
- Mean = 4
- Variance = $[(2-4)^2 + (4-4)^2 + (6-4)^2]/3 = 2.67$

- Higher variance = more spread out data.

Standard Deviation (SD)

- Definition:
- The standard deviation is the square root of the variance.
- It tells how much data values deviate from the mean.

- Formula:
- SD = √Variance

- Example:
- If Variance = 2.67 → SD = √2.67 = 1.63

- Interpretation:
- • Small SD → values close to mean.
- • Large SD → values are more spread out.

Worked Example

- Data: 5, 10, 15
- Mean = (5 + 10 + 15)/3 = 10
- Median = 10
- Mode = None
- Range = 15 - 5 = 10
- Variance = $[(5-10)^2 + (10-10)^2 + (15-10)^2]/3 = 16.67$
- SD = $\sqrt{16.67} = 4.08$

Why Descriptive Statistics are Important

- Descriptive statistics are used to:
- • Summarize large data sets into simple numbers.
- • Identify patterns and trends.
- • Compare groups of data.
- • Provide the foundation for inferential statistics.

Common Mistakes to Avoid

- • Using mean with skewed data.
- • Forgetting to arrange data before finding the median.
- • Ignoring outliers.
- • Mixing up variance and standard deviation.

Summary Table

- Measure | Purpose | Example
- Mean – Average value – 30
- Median – Middle value – 25
- Mode – Most frequent – 20
- Range – Max - Min – 40
- Variance – Spread – 16.7
- SD – Deviation from mean – 4.1

Class Activity

- Calculate the following for data set: 12, 15, 18, 20, 25
- 1. Mean
- 2. Median
- 3. Mode
- 4. Range
- 5. Variance
- 6. Standard Deviation

Conclusion

- Descriptive statistics are the foundation of data analysis.
- They help summarize, compare, and understand data clearly.
- Understanding these basic concepts is essential for all research fields.