1 Fundamentals of Statistics

- The word **Statistics** is used in both singular and plural forms.

- In singular form, it refers to the *science of collection, presentation, analysis, and interpretation of numerical data.*

- In plural form, it refers to *numerical facts collected with a definite purpose.*

—

2 Frequency and Frequency Distribution

Frequency: The number of times an observation occurs in the given data is called the frequency.
   Types of Frequency Distribution:

1. **Discrete Frequency Distribution:** Data is presented as individual observations.

2. **Continuous (Grouped) Frequency Distribution:** Data is grouped into class intervals.

Class Mark Formula:

$$\text{Class Mark} = \frac{\text{Lower Limit} + \text{Upper Limit}}{2}$$

—

3 Measures of Central Tendency

The commonly used measures of central tendency include:

- Arithmetic Mean (Mean)

- Median

- Mode

3.1 Mean of Grouped Data

Let $x_1, x_2, x_3, \ldots, x_n$ be observations with respective frequencies $f_1, f_2, f_3, \ldots, f_n$.

$$\text{Mean}(\bar{x}) = \frac{\sum f_i x_i}{\sum f_i}$$

3.1.1 Methods to Calculate Mean

There are three methods to compute the mean of grouped data:
   1. Direct Method:

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

2. Assumed Mean Method:

$$\bar{x} = a + \frac{\sum f_i d_i}{\sum f_i}$$

where:

$$d_i = x_i - a$$

a is any arbitrary value chosen as the assumed mean.
   3. Step-Deviation Method:

$$\bar{x} = a + \frac{\sum f_i u_i}{\sum f_i} \cdot h$$

where:

$$u_i = \frac{x_i - a}{h}$$

h is the class width.

—

4 Median of Grouped Data

Condition I: When the data is discrete, follow these steps:

1. Arrange the data in ascending order.

2. If the total frequency n is odd:

$$\text{Median} = \left(\frac{n+1}{2}\right)^{\text{th}} \text{ observation}$$

3. If the total frequency n is even:

$$\text{Median} = \frac{\left(\frac{n}{2}\right)^{\text{th}} \text{ observation} + \left(\frac{n}{2} + 1\right)^{\text{th}} \text{ observation}}{2}$$

Condition II: For grouped frequency distribution, the formula used is:

$$\text{Median} = l + \left(\frac{\frac{n}{2} - c}{f}\right) \cdot h$$

where:

- l = lower limit of median class

- n = total number of observations

- c = cumulative frequency of the class before median class

- f = frequency of the median class

- h = class width

—

5 Mode of Grouped Data

The class with the maximum frequency is called the **modal class**.

$$\text{Mode} = l + \left(\frac{f_1 - f_0}{2f_1 - f_0 - f_2}\right) \cdot h$$

where:

- l = lower limit of the modal class

- f_1 = frequency of the modal class

- f_0 = frequency of the class before modal class

- f_2 = frequency of the class after modal class

- h = class width

—

6 Relation Between Mean, Median, and Mode

The empirical relationship between the three measures of central tendency is:

$$\text{Mode} = 3(\text{Median}) - 2(\text{Mean})$$

—

7 Conclusion

- Statistics deals with the collection, analysis, interpretation, and presentation of data.

- Measures of central tendency help determine representative values of data.

- Different methods such as direct, assumed mean, and step-deviation methods are used to compute mean.

- Median is the middle value, and mode is the most frequently occurring value.