

# Measures of Central Tendencies

## Data Analysis

**Statistical  
Investigation**

**5. Analysis  
of data**

**A. Average**

**B.  
Dispersion**

**C.  
Correlation**

# Measures of Central Tendencies

- It is a typical value of the entire data
- It describes the characteristics of the entire group of observations
- A typical value around which other figures assemble
- This value is lies between the minimum and maximum values and generally located in the centre of the distribution

# Central Tendencies

```
graph TD; A[Central Tendencies] --> B[Mathematical Average]; A --> C[Positional Average]; B --> D[Arithmetic Mean (X)]; C --> E[Median (Md)]; C --> F[Mode (Mo)];
```

**Mathematical Average**

**Arithmetic Mean (X)**

**Positional Average**

**Median (Md)**

**Mode (Mo)**

# Arithmetic Mean

- It is simply called mean
- It is the most common type and widely used
- It is defined as the quantity obtained by adding together all the given items and by dividing this by the total number of items and is denoted as  $\bar{X}$

$$\bar{X} = \frac{\text{Sum of values of items}}{\text{Number of items}}$$

# Arithmetic Mean

- Ungrouped data
- Grouped data
  - Direct method
  - Indirect method/ Short cut method

**Merits – well defined, easy, all items considered, basic for further calculations**

**Demerits – some times not correct when values are very big or small, may give false conclusion, may give meaningless values**

# Mean for Ungrouped data

## 1. Direct method –

- Add up all the values of the variables  $x$  and get  $\Sigma x$
- Divide  $\Sigma x$  by their number of observations  $N$

$$\bar{X} = \frac{\text{Sum of values of items}}{\text{Number of items}}$$

# Mean for Ungrouped data

## 2. Short cut method –

- Assume any one value as an assumed mean A (arbitrary average)
- Find out the difference of each value from assumed mean  $d = x - A$
- Add all the  $\Sigma d$
- Apply the formula

$$\bar{X} = A \pm \frac{\Sigma d}{N}$$



# Mean for Grouped data/ Discrete series

## 1. Direct method –

- Multiply each variable by its frequency (fx)
- Add all the fx ( $\Sigma fx$ )
- Divide  $\Sigma fx$  by total number of frequency N/  $\Sigma f$

$$\bar{X} = \frac{\Sigma fx}{\Sigma f}$$

# Mean for Grouped data/ Discrete series

## 2. Short cut method –

- Assume any one value as an assumed mean A (arbitrary average)
- Find out the difference of each value from assumed mean  $d = x - A$
- Multiply each deviation by its frequency (fd)
- Add all the products ( $\Sigma fd$ )
- Apply the formula

$$\bar{X} = A \pm \frac{\Sigma fd}{\Sigma f}$$

# Mean for Grouped data/ Continuous series

## 1. Direct method –

- Find out the mid value of each class ( $10+20/2 = 15$ )
- Multiply the mid value of each class by the frequency of the class (mid x f)
- Add all the products ( $\Sigma f \text{ mid } x$ )
- Divide  $\Sigma f \text{ mid } x$  by total number of frequency  $\Sigma f$

$$\bar{X} = \frac{\Sigma f x}{\Sigma f}$$

# Mean for Grouped data/ Continuous series

## 2. Short cut method –

- Find out the mid value of each class
- Assume any one of the mid value as an assumed mean A (arbitrary average)
- Find out the difference of each value from assumed mean d = mid x – A
- Multiply each deviation by its frequency (fd)
- Add all the products ( $\Sigma fd$ )
- Apply the formula 
$$= A \pm \frac{\Sigma fd}{\Sigma f}$$

# Median

- It is an average which divides a distribution into 2 equal halves
- When the given values are arranged in ascending or descending order, the value which is in centre is the median
- It is represented by the letter Md
- When the values are odd number of items  
 $Md = N+1/2^{\text{th}}$  item
- When the values are even number of items  
 $Md = N+1/2^{\text{th}}$  item
- Median of the grouped data with class intervals are calculated as

L = Lower limit of median class

N = Total frequency

Cf = cumulative frequency

C = Class interval of median class

f = Frequency of median class

$$Md = L + \frac{\frac{N}{2} - cf}{f} C$$

# Median

- **Merits– Simple, can be calculated without knowing all the values, can also be calculated graphically**
- **Demerits – Not based on all the items, can not be used for further calculations**

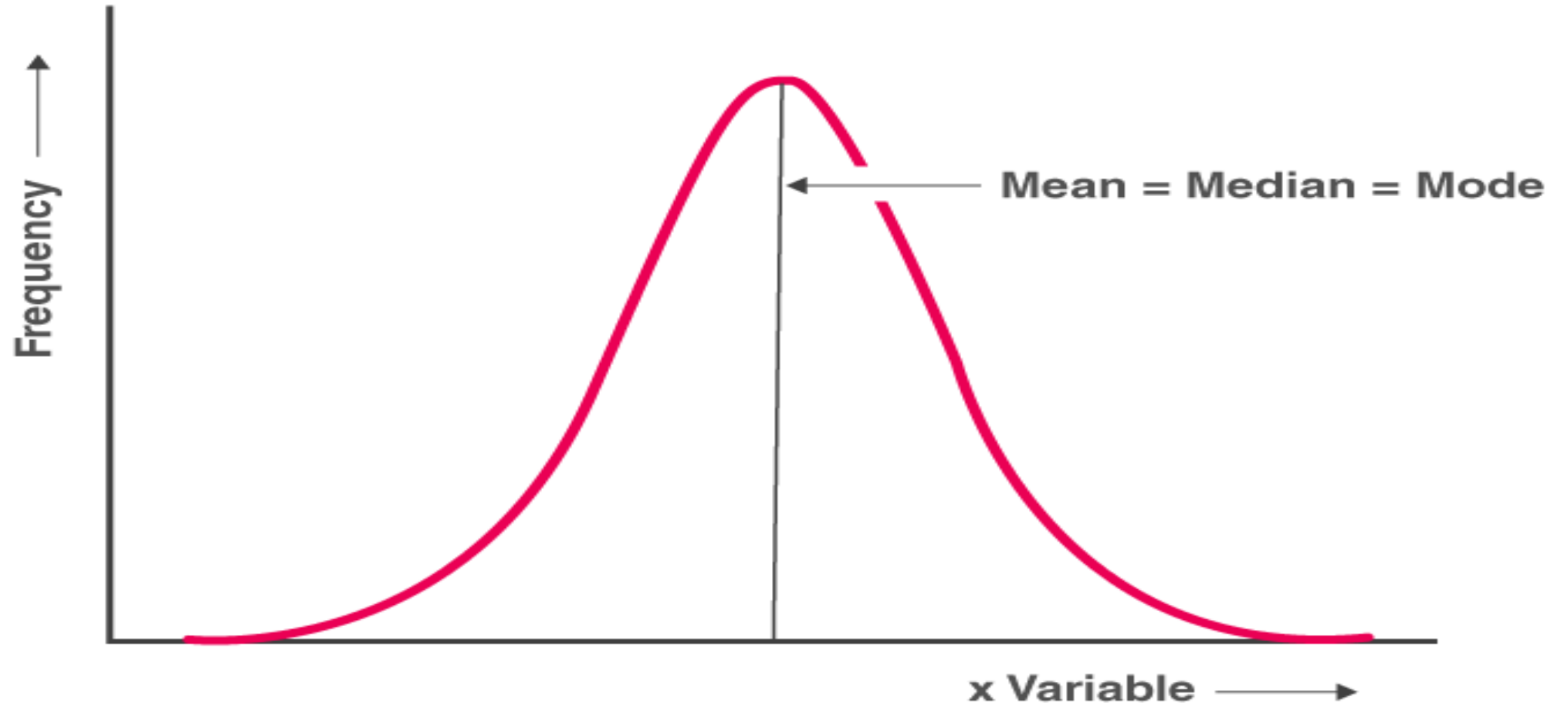
# Mode

- It is the most common item of the series
- It is defined as value of the variable which occur most frequently in the distribution
- A distribution may be unimodal or bimodal, trimodal etc.
- It is represented by  $M_o$

**Merits – Easy to find, not affected by extreme values, can be determined by graphical method**

**Demerits – Not based on all the observations, not reliable, not used for further calculations**

# An ideal Observation





# Possible Observations

