

MEASURES CENTRAL TENDENCY OR AVERAGES

If the average tends to lie at the centre of the distribution then it is called measures of central tendency or measures of location.

• Arithmetic Mean / Mean (\bar{X})

Sum of the values divided by their numbers is known as arithmetic mean. It is denoted by \bar{X} .i.e.

Methods of Arithmetic Mean

1. Direct Method
2. Short Cut Method /Indirect Method
3. Coding Method or Step-Deviation Method

1. Direct Method

For Ungrouped Data

$$\text{A.M} = \bar{X} = \frac{\sum X}{n}$$

Where n = number of observation

For Grouped Data

$$\text{A.M} = \bar{X} = \frac{\sum fx}{\sum f}$$

Where f = frequency

2. Short Cut Method

For Ungrouped Data

$$\bar{X} = A + \frac{\sum D}{n}$$

Where A =Constant,

$$D = X - A$$

For Grouped Data

$$\bar{X} = A + \frac{\sum fD}{\sum f}$$

3. Coding Method or Step-Deviation Method

For Ungrouped Data

$$\bar{X} = A + \frac{\sum u}{n} \times h$$

Where $u = \frac{X - A}{h}$ or $\frac{D}{h}$

For Grouped Data

$$\bar{X} = A + \frac{\sum fu}{\sum f} \times h$$

- **Weighted Arithmetic Mean (\bar{X}_w)**

If $X_1, X_2, X_3, \dots, X_n$ be n values with weights $W_1, W_2, W_3, \dots, W_n$ respectively then weighted average (\bar{X}_w) is defined as:

$$\bar{X}_w = \frac{X_1W_1 + X_2W_2 + X_3W_3 + \dots + X_nW_n}{W_1 + W_2 + W_3 + \dots + W_n}$$

Or

$$\bar{X}_w = \frac{\sum WX}{\sum W}$$

Where W= weight

- **Combined Mean / Grand Mean ($\bar{\bar{X}}$)**

$$\bar{\bar{X}} = \frac{\sum n\bar{X}_i}{\sum n}$$

- **Geometric Mean (G.M)**

The positive nth root of the product of n positive values is called geometric mean. It is denoted by G.M. i.e.

For Ungrouped Data

$$G.M = n\sqrt{X_1 \cdot X_2 \cdot X_3 \cdot \dots \cdot X_n}$$

Or

$$G.M = \text{Anti log} \frac{\sum \log X}{n}$$

For Grouped Data

$$G.M = \text{Anti log} \frac{\sum f \log x}{\sum f}$$

- **Weighted Geometric Mean ($G.M_w$)**

$$G.M_w = \text{Anti log} \frac{\sum W \log x}{\sum W}$$

- **Harmonic Mean (H.M)**

Harmonic mean is defined as the reciprocal of the arithmetic mean of the reciprocal of values in a data. It is denoted by H.M. i.e.

For Ungrouped Data

$$H.M = \frac{n}{\sum \frac{1}{x}}$$

For Grouped Data

$$H.M = \frac{\sum f}{\sum f \frac{1}{x}}$$

- **Median (\tilde{X})**

The central value of an array is called median. It is denoted by \tilde{X} .

For Ungrouped Data

$$\text{Median} = \tilde{X} = \text{Value of } \left[\frac{n+1}{2} \right] \text{th terms}$$

For Grouped Data

$$\text{Median} = \tilde{X} = l + \frac{h}{f} \left(\frac{\sum f}{2} - C.F \right)$$

Where

l = lower class boundary

C.F = Cumulative Frequency

h = Class interval

f = Frequency

- **Other Positional Measures**

1. Quartile
2. Deciles
3. Percentiles

1. **Quartile**

These are the values which divide a distribution into four equal parts.

$$Q_1, Q_2, Q_3$$

For Ungrouped Data

$$\text{Lower Quartile} = Q_1 = \text{Value of } \left[\frac{n+1}{4} \right] \text{ th terms}$$

$$\text{Median} = Q_2 = \text{Value of } \left[\frac{2(n+1)}{4} \right] \text{ th terms}$$

$$\text{Upper Quartile} = Q_3 = \text{Value of } \left[\frac{3(n+1)}{4} \right] \text{ th terms}$$

For Grouped Data

$$\text{Lower Quartile} = Q_1 = l + \frac{h}{f} \left(\frac{\sum f}{4} - C.F \right)$$

$$\text{Median} = Q_2 = l + \frac{h}{f} \left(\frac{2\sum f}{4} - C.F \right)$$

$$\text{Upper Quartile} = Q_3 = l + \frac{h}{f} \left(\frac{3\sum f}{4} - C.F \right)$$

Where

l = lower class boundary

C.F = Cumulative Frequency

h = Class interval

f = Frequency

2. **Deciles**

These are the values which divide a distribution into ten equal parts.

$$D_1, D_2, D_3, \dots, D_9$$

For Ungrouped Data

$$D_1 = \text{Value of } \left[\frac{n+1}{10} \right] \text{ th terms}$$

$$D_2 = \text{Value of } \left[\frac{2(n+1)}{10} \right] \text{ th terms}$$

$$D_3 = \text{Value of } \left[\frac{3(n+1)}{10} \right] \text{ th terms}$$

.....

.....

.....

$$D_8 = \text{Value of } \left[\frac{8(n+1)}{10} \right] \text{ th terms}$$

$$D_9 = \text{Value of } \left[\frac{9(n+1)}{10} \right] \text{ th terms}$$

For Grouped Data

$$D_1 = l + \frac{h}{f} \left(\frac{\sum f}{10} - C.F \right)$$

$$D_2 = l + \frac{h}{f} \left(\frac{2\sum f}{10} - C.F \right)$$

$$D_3 = l + \frac{h}{f} \left(\frac{3\sum f}{10} - C.F \right)$$

.....

.....

.....

$$D_8 = l + \frac{h}{f} \left(\frac{8\sum f}{10} - C.F \right)$$

$$D_9 = l + \frac{h}{f} \left(\frac{9\sum f}{10} - C.F \right)$$

Where

l = lower class boundary

C.F = Cumulative Frequency

h = Class interval

f = Frequency

3. Percentiles

These are the values which divide a distribution into hundred equal parts.

$$P_1, P_2, P_3, \dots, P_{99}$$

For Ungrouped Data

$$P_1 = \text{Value of } \left[\frac{n+1}{100} \right] \text{ th terms}$$

$$P_2 = \text{Value of } \left[\frac{2(n+1)}{100} \right] \text{ th terms}$$

$$P_3 = \text{Value of } \left[\frac{3(n+1)}{100} \right] \text{ th terms}$$

.....

$$\dots \dots \dots$$

$$\dots \dots \dots$$

$$P_{98} = \text{Value of } \left[\frac{98(n+1)}{100} \right] \text{ th terms}$$

$$P_{99} = \text{Value of } \left[\frac{99(n+1)}{100} \right] \text{ th terms}$$

For Grouped Data

$$P_1 = l + \frac{h}{f} \left(\frac{\sum f}{100} - C.F \right)$$

$$P_2 = l + \frac{h}{f} \left(\frac{2\sum f}{100} - C.F \right)$$

$$P_3 = l + \frac{h}{f} \left(\frac{3\sum f}{100} - C.F \right)$$

$$\dots \dots \dots$$

$$\dots \dots \dots$$

$$P_{50} = l + \frac{h}{f} \left(\frac{50\sum f}{100} - C.F \right)$$

$$\dots \dots \dots$$

$$\dots \dots \dots$$

$$P_{98} = l + \frac{h}{f} \left(\frac{98\sum f}{100} - C.F \right)$$

$$P_{99} = l + \frac{h}{f} \left(\frac{99\sum f}{100} - C.F \right)$$

Where
l = lower class boundary
 C.F = Cumulative Frequency
h = Class interval
f = Frequency

- **Mode (\hat{X})**
 "The mode which is defined as the value which occurs the greatest number of times in the data" some time, there are more than one mode in the data, and sometimes there are no mode in the data because each value occurs the same number of times. It is denoted by \hat{X} .

For Ungrouped Data

Mode = \hat{X} = Greatest value in data

For Grouped Data

$$\text{Mode} = \hat{X} = l + \frac{f_1 - f_2}{(f_1 - f_2) + (f_1 - f_3)} \times h$$

Or

$$\text{Mode} = \hat{X} = \frac{f_1 - f_2}{(2f_1 - f_2 - f_3)} \times h$$

Where

l = lower class boundary

f_1 = Frequency of the modal class or Maximum frequency

f_2 = Frequency preceding f_1

f_3 = Frequency succeeding f_1

h = Class interval of the modal class

- **Relation Between A.M , G.M and H.M**

$$A.M \geq G.M \geq H.M$$

- **Empirical relation b/w Mean , Median and Mode**

Mode = 3 Median - 2Mean

Or

$$\hat{X} = 3\tilde{X} - 2\bar{X}$$