# MEASURES OF DISPERSION, MOMENTS AND SKEWNESS

A quantity that measures how the data are dispersed about the average is called measures of dispersion.

- **Range (R)**

The range is a simplest measure of dispersion. "It is defined as the difference b/w the largest and smallest observation in a set of data." It is denoted by "R". This is an absolute measure of dispersion.

**For Ungrouped Data**

Range = R = $X_m - X_o$

Where $X_m$ = the largest value.

$X_o$ = the smallest value.

**For Grouped Data**

Range = R = Upper class boundary of the highest class – lower class boundary of the lowest class

Or

Range = R = Class Marks (X) of the highest class – Class Marks of the lowest class

- **Semi Inter Quartile Range or Quartile Deviation**

The semi inter-quartile range or quartile deviation is defined as half of the difference b/w the third and the first quartiles. Symbolically it is given by the

$$S.I.Q.R = Q.D = \frac{Q_3 - Q_1}{2}$$

Where $Q_1$ = First, Lower quartile

$Q_3$ = Third, Upper quartile

This is an absolute measure of dispersion.

- **Mean Deviation or Average Deviation**

The mean deviation is defined as the average of the deviation of the values from an average (Mean, Median), the deviation are taken without considering algebraic signs.

1. **Mean Deviation From Mean**

**For Ungrouped Data**

$$M.D = \frac{\sum\left|X - \bar{X}\right|}{n}$$

Or

$$\text{M.D} = \frac{\sum |X - Mean|}{n}$$

**For Grouped Data**

$$\text{M.D} = \frac{\sum f |X - \bar{X}|}{\sum f}$$

Or

$$\text{M.D} = \frac{\sum f |X - Mean|}{\sum f}$$

2. **Mean Deviation From Median**

**For Ungrouped Data**

$$\text{M.D} = \frac{\sum |X - \tilde{X}|}{n}$$

Or

$$\text{M.D} = \frac{\sum |X - Median|}{n}$$

**For Grouped Data**

$$\text{M.D} = \frac{\sum f |X - \tilde{X}|}{\sum f}$$

Or

$$\text{M.D} = \frac{\sum f |X - Median|}{\sum f}$$

- ## Standard Deviation (S)

The standard deviation is defined as the positive square root of the mean of the squared deviation of the values from their mean. Thus the standard deviation of a set of n values $X_1.X_2.X_3..........X_n$ .it is denoted by 'S'. This is an absolute measure of dispersion.

**Methods of Standard Deviation**

I. Direct Method
II. Short Cut Method
III. Coding Method or Step-Deviation Method

1. **Direct Method**

**For Ungrouped Data**

$$\text{S.D} = S = \sqrt{\frac{\sum X^2}{n} - \left(\frac{\sum X}{n}\right)^2}$$

$$\text{S.D} = S = \sqrt{\frac{\sum \left(X - \bar{X}\right)^2}{n}}$$

**For Grouped Data**

$$\text{S.D} = S = \sqrt{\frac{\sum fX^2}{\sum f} - \left(\frac{\sum fX}{\sum f}\right)^2}$$

$$\text{S.D} = S = \sqrt{\frac{\sum f\left(X - \bar{X}\right)^2}{\sum f}}$$

**2. Short Cut Method**

**For Ungrouped Data**

$$\text{S.D} = S = \sqrt{\frac{\sum D^2}{n} - \left(\frac{\sum D}{n}\right)^2}$$ Where D= X – A

**For Grouped Data**

$$\text{S.D} = S = \sqrt{\frac{\sum fD^2}{\sum f} - \left(\frac{\sum fD}{\sum f}\right)^2}$$

**3. Coding Method or Step-Deviation Method**

**For Ungrouped Data**

$$\text{S.D} = S = h \times \sqrt{\frac{\sum u^2}{n} - \left(\frac{\sum u}{n}\right)^2}$$ Where $u = \frac{X - A}{h} \quad or \quad \frac{D}{h}$

**For Grouped Data**

$$\text{S.D} = S = h \times \sqrt{\frac{\sum fu^2}{\sum f} - \left(\frac{\sum fu}{\sum f}\right)^2}$$

- **Combined Standard Deviation ( $S_c$ )**

**For two set of values**

$$S_c = \sqrt{\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2} + \frac{n_1 n_2}{(n_1 + n_2)^2}\left(\bar{X}_1 - \bar{X}_2\right)^2}$$

**For three or more sets of data**

$$S_c = \sqrt{\frac{\sum n_i\left[S_i^2\left(\bar{X}_i - \bar{\bar{X}}\right)^2\right]}{\sum n_i}}$$

- **Variance ( $S^2$ )**

The variance is defined as the mean of the squared deviation from mean. It is denoted by '$S^2$'

Or

The square of the standard c=deviation is called variance. It is denoted by '$S^2$'

**<u>Methods of Standard Deviation</u>**

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

**1. Direct Method**

**For Ungrouped Data**

$$\text{Var(X)} = S^2 = \frac{\sum X^2}{n} - \left(\frac{\sum X}{n}\right)^2$$

$$\text{Var(X)} = S^2 = \frac{\sum\left(X - \bar{X}\right)^2}{n}$$

**For Grouped Data**

$$\text{Var(X)} = S^2 = \frac{\sum fX^2}{\sum f} - \left(\frac{\sum fX}{\sum f}\right)^2$$

$$\text{Var(X)} = S^2 = \frac{\sum f\left(X - \bar{X}\right)^2}{\sum f}$$

**2. Short Cut Method**

**For Ungrouped Data**

$$\text{Var(X)} = S^2 = \frac{\sum D^2}{n} - \left(\frac{\sum D}{n}\right)^2$$

Where D= X – A

**For Grouped Data**

$$\text{Var(X)} = S^2 = \frac{\sum fD^2}{\sum f} - \left(\frac{\sum fD}{\sum f}\right)^2$$

**3. Coding Method or Step-Deviation Method**

**For Ungrouped Data**

$$\text{Var(X)} = S^2 = h^2 \times \left[\frac{\sum u^2}{n} - \left(\frac{\sum u}{n}\right)^2\right] \qquad \text{Where} \quad u = \frac{X-A}{h} \ \ or \ \ \frac{D}{h}$$

**For Grouped Data**

$$\text{Var(X)} = S^2 = h^2 \times \left[\frac{\sum fu^2}{\sum f} - \left(\frac{\sum fu}{\sum f}\right)^2\right]$$

- ## Combined Variance ( $S_c^{\,2}$ )

**For two set of values**

$$S_c^{\,2} = \frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2} + \frac{n_1 n_2}{\left(n_1 + n_2\right)^2}\left(\bar{X}_1 - \bar{X}_2\right)^2$$

**For three or more sets of data**

$$S_c^{\,2} = \frac{\sum n_i \left[S_i^2 \left(\bar{X}_i - \bar{\bar{X}}\right)^2\right]}{\sum n_i}$$

- ## Relative Measure of Dispersion

**1. Coefficient Of Range**

$$\text{Coefficient of Range} = \frac{X_m - X_o}{X_m + X_o}$$

**2. Coefficient Of Quartile Deviation**

$$\text{Coefficient of Q.D} = \frac{Q_3 - Q_1}{Q_3 + Q_1}$$

Where $Q_1$ = First, Lower quartile

$Q_3$ = Third, Upper quartile

**3. Coefficient Of Mean Deviation From Mean**

Coefficient of M.D from Mean = $\frac{Mean\ Deviation\ From\ Mean}{Mean}$

Or

Coefficient of M.D from Mean = $\frac{M.D\ From\ \bar{X}}{\bar{X}}$

4. **Coefficient Of Mean Deviation From Median**

Coefficient of M.D from Median = $\frac{Mean\ Deviation\ From\ Median}{Median}$

Or

Coefficient of M.D from Mean = $\frac{M.D\ From\ \tilde{X}}{\tilde{X}}$

5. **Coefficient Of Standard Deviation**

Coefficient of S.D = $\frac{S.D}{\bar{X}}$

6. **Coefficient Of Variation (C.V)**

"The coefficient of variation expresses the standard deviation as a percentage in terms of arithmetic mean". It is used as a criterion of consistent performance, the smaller coefficient of variation, and the more consistent in the performance.

Or

"Coefficient of variation is used to compare the variability of two or more than two series".

Coefficient of Variation = C.V = $\frac{S.D}{\bar{X}} \times 100$

- ## Relationship Between Measures of Dispersion

1. **For Normal Distribution**

I. Mean Deviation = M.D = 0.7979 S.D

II. Quartile Deviation = Q.D = 0.6745 S.D

2. **For Moderately Skewed Distribution**

I. Mean Deviation = M.D = $\frac{3}{4}$ S.D

II. Quartile Deviation = Q.D = $\frac{2}{3}$ S.D

III. Quartile Deviation = Q.D = $\frac{5}{6}$ M.D

- **Moments**

A moment designates the power to which deviation are raised before averaging them.

**Methods of Standard Deviation**

1. Moments about Mean or Central Moments
2. Moments about Origin or Zero
3. Moments about Provisional Mean or Arbitrary Value (Non Central Moment)

**1. Moments about Mean or Central Moments**

**For Ungrouped Data**

$$\mu_1 = m_1 = \frac{\sum\left(x-\bar{x}\right)}{n} = 0$$

$$\mu_2 = m_2 = \frac{\sum\left(x-\bar{x}\right)^2}{n} = Variance$$

$$\mu_3 = m_3 = \frac{\sum\left(x-\bar{x}\right)^3}{n}$$

$$\mu_4 = m_4 = \frac{\sum\left(x-\bar{x}\right)^4}{n}$$

**For Grouped Data**

$$\mu_1 = m_1 = \frac{\sum f\left(x-\bar{x}\right)}{\sum f} = 0$$

$$\mu_2 = m_2 = \frac{\sum f\left(x-\bar{x}\right)^2}{\sum f} = Variance$$

$$\mu_3 = m_3 = \frac{\sum f\left(x-\bar{x}\right)^3}{\sum f}$$

$$\mu_4 = m_4 = \frac{\sum f\left(x-\bar{x}\right)^4}{\sum f}$$

**2. Moments about Origin or Zero**

**For Ungrouped Data**

$$\mu'_1 = m'_1 = \frac{\sum x}{n}$$

$$\mu'_2 = m'_2 = \frac{\sum x^2}{n}$$

$$\mu'_3 = m'_3 = \frac{\sum x^3}{n}$$

$$\mu'_4 = m'_4 = \frac{\sum x^4}{n}$$

**For Grouped Data**

$$\mu'_1 = m'_1 = \frac{\sum fx}{\sum f}$$

$$\mu'_2 = m'_2 = \frac{\sum fx^2}{\sum f}$$

$$\mu'_3 = m'_3 = \frac{\sum fx^3}{\sum f}$$

$$\mu'_4 = m'_4 = \frac{\sum fx^4}{\sum f}$$

3. **Moments about Provisional Mean or Arbitrary Value (Non Central Moment)**
4.

**<u>Methods of Standard Deviation</u>**

i. Direct Method
ii. Short Cut Method
iii. Coding Method or Step-Deviation Method

**i. Direct Method**

**For Ungrouped Data**

$$\mu'_1 = m'_1 = \frac{\sum (x - A)}{n}$$

Where A is constant

$$\mu'_2 = m'_2 = \frac{\sum (x - A)^2}{n}$$

$$\mu'_3 = m'_3 = \frac{\sum (x - A)^3}{n}$$

$$\mu'_4 = m'_4 = \frac{\sum (x - A)^4}{n}$$

**For Grouped Data**

$$\mu'_1 = m'_1 = \frac{\sum f(x-A)}{\sum f}$$ Where A is constant

$$\mu'_2 = m'_2 = \frac{\sum f(x-A)^2}{\sum f}$$

$$\mu'_3 = m'_3 = \frac{\sum f(x-A)^3}{\sum f}$$

$$\mu'_4 = m'_4 = \frac{\sum f(x-A)^4}{\sum f}$$

**ii. Short Cut Method**

**For Ungrouped Data**

$$\mu'_1 = m'_1 = \frac{\sum D}{n}$$ Where D= X - A

$$\mu'_2 = m'_2 = \frac{\sum D^2}{n}$$

$$\mu'_3 = m'_3 = \frac{\sum D^3}{n}$$

$$\mu'_4 = m'_4 = \frac{\sum D^4}{n}$$

**For Grouped Data**

$$\mu'_1 = m'_1 = \frac{\sum fD}{\sum f}$$ Where D= X - A

$$\mu'_2 = m'_2 = \frac{\sum fD^2}{\sum f}$$

$$\mu'_3 = m'_3 = \frac{\sum fD^3}{\sum f}$$

$$\mu'_4 = m'_4 = \frac{\sum fD^4}{\sum f}$$

**iii. Coding Method or Step Deviation Method**

**For Ungrouped Data**

$$\mu'_1 = m'_1 = \frac{\sum u}{n} \times h \qquad \text{Where} \quad u = \frac{X-A}{h} \quad or \quad \frac{D}{h}$$

$$\mu'_2 = m'_2 = \frac{\sum u^2}{n} \times h^2$$

$$\mu'_3 = m'_3 = \frac{\sum u^3}{n} \times h^3$$

$$\mu'_4 = m'_4 = \frac{\sum u^4}{n} \times h^4$$

**For Grouped Data**

$$\mu'_1 = m'_1 = \frac{\sum fu}{\sum f} \times h \qquad \text{Where} \quad u = \frac{X-A}{h} \quad or \quad \frac{D}{h}$$

$$\mu'_2 = m'_2 = \frac{\sum fu^2}{\sum f} \times h^2$$

$$\mu'_3 = m'_3 = \frac{\sum fu^3}{\sum f} \times h^3$$

$$\mu'_4 = m'_4 = \frac{\sum fu^4}{\sum f} \times h^4$$

- **Relation Between Central moments in Terms of Non Central Moments**

$$\mu_1 = m_1 = \mu'_1 - \mu_1' = 0$$

$$\mu_2 = m_2 = \mu'_2 - \left(\mu_1'\right)^2 = Varaince$$

$$\mu_3 = m_3 = \mu'_3 - 3\mu_1'\mu_2' + 2\left(\mu_1'\right)^3$$

$$\mu_4 = m_4 = \mu'_4 - 4\mu_1'\mu_3' + 6\left(\mu_1'\right)^2 \mu_2' - 3\left(\mu_1'\right)^4$$

- **Moments – Ration**

$$\beta_1 = b_1 = \frac{\mu_3^{\,2}}{\mu_2^{\,3}}$$

$$\beta_2 = b_2 = \frac{\mu_4}{\mu_2^{\,2}}$$

- **Sheppard's Correction for Moments of Group Data**

$$\mu_2(corrected) = \mu_2(uncorrected) - \frac{h^2}{12}$$

$$\mu_3(corrected) = \mu_3$$

$$\mu_4(corrected) = \mu_4(uncorrected) - \frac{h^2}{2}\mu_2(uncorrected) + \frac{7}{240}h^4$$

- **Charliers Check**

i. $\sum f(u+1) = \sum fu + \sum f$

ii. $\sum f(u+1)^2 = \sum fu^2 + 2\sum fu + \sum f$

iii. $\sum f(u+1)^3 = \sum fu^3 + 3\sum fu^2 + 3\sum fu + \sum f$

iv. $\sum f(u+1)^4 = \sum fu^4 + 4\sum fu^3 + 6\sum fu^2 + 4\sum fu + \sum f$

- **Symmetry**
In a symmetrical distribution a deviation below the mean exactly equals the corresponding deviation above the mean. It is called symmetry.
For symmetrical distribution the following relations hold.

Mean = Median = Mode

$Q_3$ - Median = Median - $Q_1$

$u_3 = m_3 = 0$

$\beta_1 = b_1 = 0$

- **Skewness**
Skewness is the lack of symmetry in a distribution around some central value i.e. means Median or Mode. It is the degree of asymmetry.

Mean $\neq$ Median $\neq$ Mode

$Q_3$ - Median $\neq$ Median - $Q_1$

$u_3 = m_3 \neq 0$

$\beta_1 = b_1 \neq 0$

There are two types of Skewness.

1. **Positive Skewness**
   If the frequency curve has a longer tail to right, the distribution is said to be positively skewed.
2. **Negative Skewness**
   If the frequency curve has a longer tail to left, the distribution is said to negatively skewed.

- ## Coefficient of Skewness (SK)

**Karl Pearson's Coefficient of Skewness**

**SK =** $\frac{Mean - Mode}{S.D}$

**SK =** $\frac{3(Mean - Median)}{S.D}$

**Bowly's Quartile Coefficient of Skewness**

**SK =** $\frac{Q_3 + Q_1 - 2Median}{Q_3 - Q_1}$

**Moment Coefficient of Skewness**

**SK =** $\frac{\sqrt{\beta_1}(\beta_2 + 3)}{2(5\beta_2 - 6\beta_1 - 9)}$

- ## Kurtosis

Moment coefficient $\beta_2$ is an important measure of kurtosis. These measures define as:

$$\beta_2 = b_2 = \frac{\mu_4}{\mu_2^{\ 2}}$$

The moment coefficient $\beta_2$ is a pure numbers and independent of the origin and unit of measurement.

If $\beta_2 > 3$ distribution is Leptokurtic
If $\beta_2 = 3$ distribution is Normal or Mesokurtic
If $\beta_2 < 3$ distribution is Platy Kurtic
Or
$K = \frac{Q.D}{P_{90} - P_{10}}$ For Normal distribution, K = 0.263