❖ **Moments: "The arithmetic mean of the r^{th} power of deviations taken either from mean, zero or from any arbitrary origin (provisional ,mean) are called moments".**

- When the deviations are computed from the arithmetic mean, then such moments are called moments about mean (mean moments) or sometimes called **central moments**, denoted by m_r and given as follows:

	Ungrouped Data	Grouped Data
For Sample	$m_r = \dfrac{\sum (x_i - \bar{x})^r}{n}$	$m_r = \dfrac{\sum f(x_i - \bar{x})^r}{n}$
For Population	$\mu_r = \dfrac{\sum (x_i - \mu)^r}{N}$	$\mu_r = \dfrac{\sum f(x_i - \mu)^r}{N}$
Where r = 1,2,3,4…		

- When the deviations of the values are computed from origin or zero, then such moments ar called the **moments about origin**, denoted by m'_r and are given by:

	Ungrouped Data	Grouped Data
For Sample	$m'_r = \dfrac{\sum x_i^r}{n}$	$m'_r = \dfrac{\sum f x_i^r}{n}$
For Population	$\mu'_r = \dfrac{\sum x_i^r}{N}$	$\mu'_r = \dfrac{\sum f x_i^r}{N}$
Where r = 1,2,3,4…		

- When the deviations of the values are computed from any arbitrary value say A (provisional mean), then such moments are called moments about provisional mean, denoted by m'_r or μ'_r. There are two methods for finding moments about provisional mean: i.e. Short-cut Method and Step-deviation Method.

	Ungrouped Data		Grouped Data	
	Short cut Method	Step-deviation Method	Short cut Method	Step-deviation Method
For Sample	$m'_r = \frac{\sum D^r}{n}$	$m'_r = h^r \left(\frac{\sum ui^r}{n} \right)$	$m'_r = \frac{\sum fD^r}{n}$	$m'_r = h^r \left(\frac{\sum fui^r}{n} \right)$
For Population	$\mu'_r = \frac{\sum D^r}{N}$	$\mu'_r = h^r \left(\frac{\sum ui^r}{N} \right)$	$\mu'_r = \frac{\sum fD^r}{N}$	$\mu'_r = h^r \left(\frac{\sum fui^r}{N} \right)$
Where r = 1,2,3,4… D= x – A and where h is the common width of the class intervals, $ui = \frac{xi - A}{h}$				

All the raw moments can then be converted into central moments or mean moments or moments about mean, by using the following relations:

For Sample:

$$m_1 = 0$$
$$m_2 = m'_2 - (m'_1)^2$$
$$m_3 = m'_3 - 3m'_1 m'_2 + 2(m'_1)^3$$
$$m_4 = m'_4 - 4m'_1 m'_3 + 6(m'_1)^2 m'_2 - 3(m'_1)^4$$

For Population:

$$\mu_1 = 0$$
$$\mu_2 = \mu'_2 - (\mu'_1)^2$$
$$\mu_3 = \mu'_3 - 3\mu'_1 \mu'_2 + 2(\mu'_1)^3$$
$$\mu_4 = \mu'_4 - 4\mu'_1 \mu'_3 + 6(\mu'_1)^2 \mu'_2 - 3(\mu'_1)^4$$

- Numerical example of first Moments:

- Calculate first four moments about mean for ungrouped data for the following set of examination marks:

General formula for moment about mean are given below:

$m_r = \frac{\sum (x_i - \bar{x})^r}{n}$, where *r=1, 2, 3, 4.*

Put r=1 $m_1 = \frac{\sum (x_i - \bar{x})}{n}$, put r=2 $m_2 = \frac{\sum (x_i - \bar{x})^2}{n}$, put r=3

$m_3 = \frac{\sum (x_i - \bar{x})^3}{n}$ and put r=4

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

x_i	$x_i-\bar{x}$	$(x_i-\bar{x})^2$	$(x_i-\bar{x})^3$	$(x_i-\bar{x})^4$
32	-8	64	-512	4096
36	-4	16	-64	256
36	-4	16	-64	256
37	-3	09	-27	81
39	-1	01	-1	1
41	1	01	1	1
45	5	25	125	625
46	6	36	216	1296
48	8	64	512	4096
$\sum x_i=360$	$\sum(x_i-\bar{x}) = 0$	$\sum(x_i-\bar{x})^2 =232$	$\sum(x_i-\bar{x})^3 = 186$	$\sum(x_i-\bar{x})^4 = 10708$

$m_1 = 0$, $m_2 = 25.78$ (marks)^2, $m_3 = 20.67$ (marks)^3 , $m_4 = 1189.78$ (marks)^4 (Answer).

❖ Question # Calculate first four moments about mean for grouped data (using a continuous grouped case formula). The following distribution relates to the number of assistants in 50 retail establishments, the data are given below:

No. of assistants	0	1	2	3	4	5	6	7	8	9
f	3	4	6	7	10	6	5	5	3	1

Using these formulae

$$m_1=\frac{\sum f_i(x_i-\bar{x})}{\sum f_i}\ ,\quad m_2=\frac{\sum f_i(x_i-\bar{x})^2}{\sum f_i}\ ,\quad m_3=\frac{\sum f_i(x_i-\bar{x})^3}{\sum f_i}\quad \text{and}$$

$$m_4=\frac{\sum f_i(x_i-\bar{x})^4}{\sum f_i}$$

❖ Numerical example of Moment in continuous grouped data:

❖ Compute the first four moments and measure of Skewness and Kurtosis for the following distribution of wages using a short cut method:

Weekly earnings (Rupees)	5	6	7	8	9	10	11	12	13	1 4	15
No. of men	1	2	5	10	20	51	22	11	5	3	1

Earnings in Rs. (x_i)	Men f_i	$D_i = (x_i - A)$ $A = 10$	$f_i D_i$	$f_i D_i^2$	$f_i D_i^3$	$f_i D_i^4$
5	1	-5	-5	25	-125	625
6	2	-4	-8	32	-128	512
7	5	-3	-15	45	-135	405
8	10	-2	-20	40	-80	160
9	20	-1	-20	20	-20	20
10	51	0	0	0	0	0
11	22	1	22	22	22	22
12	11	2	22	44	88	176
13	5	3	15	45	135	405
14	3	4	12	48	192	768
15	1	5	5	25	125	625
	$\sum f_i = 131$		$\sum f_i D_i = 8$	$\sum f_i D_i^2 = 346$	$\sum f_i D_i^3 = 74$	$\sum f_i D_i^4 = 3718$

$m_1' = \dfrac{\sum f_i D_i}{\sum f_i} \Rightarrow m_1' = \dfrac{8}{131} \Rightarrow m_1' = 0.06$, $\quad m_2' = \dfrac{\sum f_i D_i^2}{\sum f_i} \Rightarrow m_2' = \dfrac{346}{131} \Rightarrow m_2' =$ 2.64

$m_3' = \dfrac{\sum f_i D_i^3}{\sum f_i} \Rightarrow m_3' = \dfrac{74}{131} \Rightarrow m_3' = 0.56$, $\quad m_4' = \dfrac{\sum f_i D_i^4}{\sum f_i} \Rightarrow m_4' = \dfrac{3718}{131}$
$\Rightarrow m_4' = 28.38$

$m_1 = m_1' - m_1' = 0$ (always zero), $\quad m_2 = m_2' - (m_1')^2 \Rightarrow m_2 = 2.64 - (0.06)^2$
$\Rightarrow m_2 = 2.64$;

$m_3 = m_3' - 3m_2' m_1' + 2(m_1')^3 \Rightarrow m_3 = 0.56 - 3(2.64)(0.06) + 2(0.06)^3 \Rightarrow m_3 = 0.08$;

$m_4 = m_4' - 4m_3' m_1' + 6m_2' (m_1')^2 - 3(m')^4$
$\Rightarrow m_4 = 28.38 - 4(0.56)(0.06) + 6(2.64)(0.06)^2 - 3(0.06)^4$

$\Rightarrow m_4 = 28.30$ (Answer).

- Calculate measure of Skewness:

$$b_1 = \frac{m_3}{m_2^3} \quad \Rightarrow b_1 = \frac{0.08}{(2.64)^2} \quad \Rightarrow b_1 = 0.0114 \qquad \text{(Answer).}$$

- Calculate measure of Kurtosis:

$$b_2 = \frac{m_4}{m_2^2} \quad \Rightarrow b_2 = \frac{28.30}{(2.64)^2} \quad \Rightarrow b_2 = 4.0604 \qquad \text{(Answer).}$$

Question # Calculate first four moment about the mean and Measure of Skewness and Kurtosis for the following data are given below,

Using a step deviation method for grouped data?

Age nearest birth day	22	27	32	37	42	47	52
No. of men	1	2	26	22	20	15	14

Symmetrical Distribution

- A distribution in which the values of mean, median and mode are equal is called symmetrical distribution i.e.

$$Mean = Median = Mode$$

- A distribution is which the two quartiles are equidistant from the median is called a symmetrical distribution i.e.

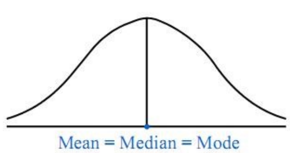
$$Q_3 + Q_1 - 2\ Median = 0$$

- A distribution is said to be symmetrical if:

$$b_1 = 0$$

- A distribution in which the two tails are equal in length from the central value then it is called symmetrical distribution. The symmetrical distribution is always in the form of a bell.

Moment-Ratios	1st Moments Ratio	2nd Moments Ratio
Sample	$b_1 = \frac{m_3^2}{m_2^3}$	$b_2 = \frac{m_4}{m_2^2}$
Population	$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$	$\beta_2 = \frac{\mu_4}{\mu_2^2}$
Moment-Ratios are independent of the origin and units of measurements i.e. they are dimensionless quantities.		



Skewness: We know that for symmetrical distribution the values of mean, median and mode are equal and that the two tails of the distribution are equal in length from the central value etc.

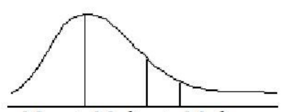
"Skewness is the degree of asymmetry" OR

"Skewness is the lack (absence) of symmetry around central value (average)"

The presence skewness tells us that a particular distribution is not symmetrical or in other words it is skewed. In skewed distribution the curve is turned more to one side than the other.

Positive Skewness

- Skewness is said to be positive, if mean is greater than the median and median is greater than mode i.e. $Mean > Median > Mode$
- Skewness is said to be positive, if: $Q_3 + Q_1 - 2\ Median > 0$
- In terms of moments, skewness is said to be positive if: $\alpha_3 > 0$
- Skewness is said to be positive, if the right tail of a distribution is longer than its left tail.

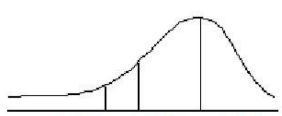


$Mean > Median > Mode$

$$\boxed{\alpha_3 = \sqrt{b_1} = \frac{m_3}{\sqrt{m_2^3}}}$$

Negative Skewness

- Skewness is said to be negative, if mean is smaller than the median and median is smaller than mode i.e. $Mean > Median > Mode$
- Skewness is said to be negative, if: $Q_3 + Q_1 - 2\ Median > 0$
- In terms of moments, skewness is said to be negative if: $\alpha_3 < 0$
- Skewness is said to be negative, if the left tail of a distribution is longer than its right tail.



$Mean < Median < Mode$

Karl Pearson's measures of Skewness: It is defined as:

$$S_k = \frac{Mean - Mode}{Standard\ Deviation}$$

It is to be noted that, this measure is suggested by Karl Pearson (1857-1936) and is known as Pearsonian coefficient.

Since in many cases mode is ill-defined, therefore we replace (Mean – Mode) by its equivalent from the empirical relation i.e. 3 (Mean – Median) and hence:

$$S_k = \frac{3(Mean - Median)}{Standard\ Deviation}$$

This coefficient usually varies between –3 and +3.

Bowley's measures of Skewness: It is defined as:

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

It is also to be noted that, this measure is suggested by Bowley (1869-1957) and is known as Bowley's coefficient.

This coefficient usually varies between –1 and +1.

Coefficient of Skewness based on Moments: It is defined by:

$$\alpha_3 = \sqrt{b_1} = \frac{m_3}{\sqrt{m_2^3}} \qquad \text{(For sample)}$$

$$\gamma_1 = \sqrt{\beta_1} = \frac{\mu_3}{\sqrt{\mu_2^3}} \qquad \text{(For population)}$$

Normal Distribution: A distribution is said to be normal if its b_1=0 and b_2=3 respectively. The curve of the normal distribution is Bell-shaped and symmetric.

For a Bell-shaped symmetric distribution:

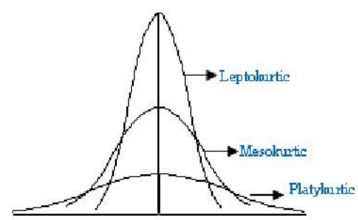
- ✓ 68.27% area of the normal curve lies under the range $\mu \pm \delta$
- ✓ 95.45% area of the normal curve lies under the range $\mu \pm 2\delta$
- ✓ 99.73% area of the normal curve lies under the range $\mu \pm 3\delta$
- ✓ Mean Deviation = 4/5 Standard Deviation
- ✓ Quartile Deviation = 2/3 Standard Deviation
- ✓ Quartile Deviation = 5/6 Mean Deviation

Kurtosis: *"The degree of peakedness or flatness of a frequency distribution relative to normal distribution is called Kurtosis". OR*

"The characteristic by which we compare the "hump" of a distribution with normal distribution is called kurtosis".

Kurtosis indicates whether a particular distribution is flatter or more peaked than the normal curve. Kurtosis is measured by the b2

- If $b_2 > 3$, then the distribution is known as **leptokurtic**
- If $b_2 = 3$, then the distribution is known as **mesokurtic**
- If $b_2 < 3$, then the distribution is known as **platykurtic**



To measure the skewness we will use: $b_2 = \frac{m_4}{m_2^2}$

Importance of Moments: It is a known fact that Statistics is divided into two types i.e. descriptive statistics and inferential statistics. Moments are the only concept in whole of the Statistics, which belong to both descriptive as well as inferential statistics. Now the point arises in what way moments are descriptive and in what way moments are inferential. To understand in what way moments are descriptive, let us consider the following table:

Moments as Descriptive Measures	
1st moment about origin	m_1' or μ_1' = Arithmetic Mean
2nd moment about mean	m_2 or μ_2 = Variance
Coefficient of skewness	b_1 or β_1 = Measure of Skewness
Coefficient of kurtosis	b_2 or β_2 = Measure of Kurtosis

So far we have discussed four characteristics of a frequency distribution given as follows:

- Measure of Central Tendency
- Measure of Dispersion
- Measure of Skewness
- Measure of Kurtosis

All of these Descriptive characteristics can be obtained by the moments.

On the other hand, Moments are included in inferential statistics in the sense that there is a very old method of estimation, commonly known as estimation by method of moments. In estimation, by the method of moments we find as many moments as the number of parameters desired to be estimated both from the theoretical distribution and from the sample data. By equating corresponding moments and solving the equations obtained in this we, we get the estimators for the parameters under consideration.

Q: Explain the term skewness and kurtosis.

Skewness: *"Skewness is the lack of symmetry around the average".*

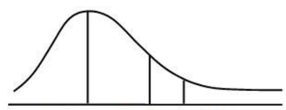
The presence of skewness tells us that a particular distribution is not symmetrical. In skewed distribution the curve is turned more to one side than the other.

There are two types of skewness:

- Positive skewness
- Negative skewness

Positive Skewness

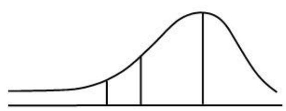
- Skewness is positive, if Mean > Median > Mode
- In terms of moments, skewness is positive if: $\sqrt{b_1}$ or $\sqrt{\beta_1} > 0$
- Skewness is positive, if the right tail of a distribution is longer than its left tail.



Mean > Median > Mode

Negative Skewness

- Skewness is negative, if Mean < Median < Mode
- In terms of moments, skewness is negative if: $\sqrt{b_1}$ or $\sqrt{\beta_1} < 0$
- Skewness is negative, if the left tail of a distribution is longer than its right tail.



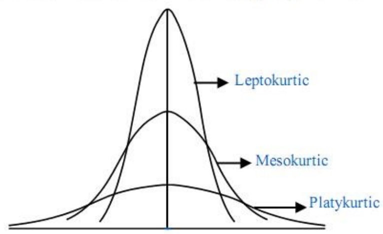
Mean < Median < Mode

Kurtosis: *"The characteristic by which we compare the "hump" of a distribution with normal distribution is called kurtosis".*

Kurtosis indicates whether a particular distribution is flatter or more peaked than the normal curve.

Kurtosis is measured by the moment ratio b_2 or β_2

- If $\beta_2 > 3$, then the distribution is known as leptokurtic
- If $\beta_2 = 3$, then the distribution is known as mesokurtic
- If $\beta_2 < 3$, then the distribution is known as platykurtic.



Q: Explain the difference between absolute and relative measure of dispersion.

Absolute Measure of Dispersion: "An absolute measure of dispersion is one that measures the dispersion in terms of the same units or in the square of units as the units of the data" e.g. if the units of the data are Rs, meters, kg, etc. The units of the measures of dispersion will also be Rs, meters, kg, etc.

The common absolute measures of dispersion are:

- Range
- Quartile Deviation
- Mean Deviation
- Standard Deviation

Relative Measure of Dispersion: "A relative measure of dispersion is one; that is expressed in the form of a ratio or percentage and independent of the units of measurements"
This is a pure number and independent of the units in which the data has been expressed. It is used for the purpose to compare the dispersion of one data set with the dispersion of another.

The common relative measures of dispersion are:

- Coefficient of Dispersion or Coefficient of Range
- Coefficient of Quartile Deviation
- Coefficient of Mean Deviation
- Coefficient of Standard Deviation or Coefficient of Variation (C.V)

Q: State the measures commonly employed to define skewness and kurtosis. What aspects of the frequency curve are measured by them?

There are various measures to define skewness and kurtosis but the following are commonly employed measures to define skewness and kurtosis:

To define skewness we use $\gamma_1 = \sqrt{\beta_1} = \frac{\mu_3}{\sqrt{\mu_2^3}}$

The following aspects of the frequency curve are measured by it:

- If $\gamma_1 = 0$ then the frequency curve is symmetric.
- If $\gamma_1 < 0$ then the frequency curve is negatively skewed i.e. the left tail of the frequency curve is longer.
- If $\gamma_1 > 0$ then the frequency curve is positively skewed i.e. the right tail of the frequency curve is longer.

It is to be noted that Karl Pearson and Bowley's coefficients can also be used to define skewness.

Similarly to define kurtosis we use $\beta_2 = \frac{\mu_4}{\mu_2^2}$

The following aspects of the frequency curve are measured by it:

- If $\beta_2 = 3$ then the frequency curve is mesokurtic i.e. it is similar to the Normal curve
- If $\beta_2 > 3$ then the frequency curve is leptokurtic i.e. it is peak as compared to the Normal curve
- If $\beta_2 < 3$ then the frequency curve is platykurtic i.e. it is flatter as compared to the Normal curve

Q: Define moment's ratios b_1 and b_2 and state the purpose for which they are used.

Moments Ratios: There are certain ratios in which both the numerator and denominators are the moments about mean. These ratios are called moment's ratios and are denoted by β_1 and β_2 or b_1 and b_2, where:

	1stMoments Ratio	2nd Moments Ratio
For Sample	$b_1 = \frac{m_3^2}{m_2^3}$	$b_2 = \frac{m_4}{m_2^2}$
For Population	$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$	$\beta_2 = \frac{\mu_4}{\mu_2^2}$

The moment ratios are independent of origin and units of measurements.

Purpose: It is to be noted that, $\sqrt{\beta_1}$ or $\sqrt{b_1}$ is the measure of skewness while β_2 or b_2 is the measure of kurtosis. Hence moment ratios are use to determined the shape of the distribution.