MOMENTS, SKEWNESS AND KURTOSIS

Introduction

In the preceding lessons, we have discussed the measures of central tendency and dispersion in case of frequency distribution. The measures of central tendency tell us about the concentration of the observations about the middle of the distribution and the measure of dispersion gives us an idea about the spread or scatter of the observations about some measure of central tendency. These measures, however, don't adequately describe a frequency distribution in the sense that there could be two or more distributions with the same mean and standard deviation but still different from each other with regard to shape or pattern of distribution. Measures of central tendency and dispersion are inadequate to characterize a distribution completely and must be supported and supplemented by two more measures viz. skewness and kurtosis which we shall discuss in this lesson.

4.1 Moments:

Moments are the general statistical measure used to describe and analyse the characteristics of a frequency distribution viz. central tendency, dispersion, skewness and kurtosis. Let us consider the variable X having a frequency distribution as given below:

X	X_1	X_2	X_3	----	---	X_n
F	f_1	f_2	f_3	----	---	f_n

Then, $\bar{x} = \dfrac{\sum f_i x_i}{\sum f_i} = \dfrac{\sum f_i x_i}{N}$ is the arithmetic mean.

4.1.2 Moments about Mean

The r^{th} moment about mean $\overline{x}$ is denoted by μ_r and also known as central moment and is defined as

$$\mu_r = \frac{\sum_{i=1}^{n} f_i \left(x_i - \overline{x}\right)^r}{N} \qquad r = 0,1,2,3.. \qquad \text{(Eq. 4.1)}$$

Putting r=0 in equation (4.1) we get

$$\mu_0 = \frac{\sum_{i=1}^{n} f_i \left(x_i - \overline{x}\right)^0}{N} = 1 \qquad \text{(Eq. 4.2)}$$

Putting r =1 in equation (4.1) we get $\mu_1 = \frac{\sum_{i=1}^{n} f_i \left(x_i - \overline{x}\right)^1}{N} = 0$

(Eq. 4.3)

because the algebraic sum of deviations of a given set of observations from their mean is zero. Thus the first moment about mean is always zero.

Again taking r =2 we get $\mu_2 = \frac{\sum_{i=1}^{n} f_i \left(x_i - \overline{x}\right)^2}{N} = \sigma^2$ (Eq. 4.4)

Hence second moment about mean gives the variance of the distribution.

$$\mu_3 = \frac{\sum_{i=1}^{n} f_i \left(x_i - \overline{x}\right)^3}{N}$$

(Eq. 4.5)

$$\mu_4 = \frac{\sum_{i=1}^{n} f_i \left(x_i - \overline{x}\right)^4}{N}$$

(Eq. 4.6)

4.1.3 Moments about an Origin

The r^{th} moment about origin denoted by μ_r', are also known as raw moment and is defined as

$$\mu_r' = \frac{1}{N}\sum_{i=1}^{n} f_i x_i^r \text{ ; } r = 0,1,2,3 \qquad \text{(Eq. 4.7)}$$

Putting r =0 and r =1 in equation we get respectively

$$\mu_0' = \frac{1}{N}\sum_{i=1}^{n} f_i x_i^0 = 1 \qquad \text{(Eq. 4.8)}$$

$$\mu_1' = \frac{1}{N}\sum_{i=1}^{n} f_i x_i^1 = \bar{x} \qquad \text{(Eq.4.9)}$$

Where, μ_1' is the first raw moment about an origin

Taking r =2, 3, 4 in (Eq. 4.7) we get,

Second moment about an origin

$$\mu_2' = \frac{1}{N}\sum_{i=1}^{n} f_i x_i^2 \qquad \text{(Eq. 4.10)}$$

Third moment about an origin $\mu_3' = \frac{1}{N}\sum_{i=1}^{n} f_i x_i^3$ (Eq. 4.11)

Fourth moment about an origin $\mu_4' = \frac{1}{N}\sum_{i=1}^{n} f_i x_i^4$ (Eq. 4.12)

4.1.4 Relation between moments about mean and moments about an origin:-

$$\mu_r = \mu_r' - {}^rc_1\mu_{r-1}'\mu_1' + {}^rc_2\mu_{r-2}'(\mu_1')^2 - {}^rc_3\mu_{r-3}'(\mu_1')^3 + \ldots + (-1)^r(\mu_1')^r \qquad \text{(Eq. 4.13)}$$

Putting r = 2, 3, 4 respectively, we get

$$\mu_2=\mu_2'-{}^2c_1\mu_1'\mu_1'+{}^2c_2\mu_0'(\mu_1')^2$$

$$\mu_2=\mu_2'-2\mu_1'\mu_1'+{}^2c_2\mu_0'(\mu_1')^2 \quad \left(\because {}^nc_n = 1 \text{ and } {}^nc_{n-1} = n\right)$$

$$\mu_2=\mu_2'-(\mu_1')^2 \qquad (\text{Eq. } 4.14)$$

$$\mu_3=\mu_3'-{}^3c_1\mu_2'+{}^3c_2\mu_1'(\mu_1')^2-{}^3c_3\mu_0'(\mu_1')^3$$

$$\mu_3=\mu_3'-3\mu_2'\mu_1'+3\mu_1'(\mu_1')^2-(\mu_1')^3 \quad \left(\because {}^nc_n = 1 \text{ and } {}^nc_{n-1} = n\right)$$

$$\mu_3=\mu_3'-3\mu_2'\mu_1'+3(\mu_1')^3-(\mu_1')^3$$

$$\mu_3=\mu_3'-3\mu_2'\mu_1'+2(\mu_1')^3 \qquad (\text{Eq. } 4.15)$$

$$\mu_4=\mu_4'-{}^4c_1\mu_3'\mu_1'+{}^4c_2\mu_2'(\mu_1')^2-{}^4c_3\mu_1'(\mu_1')^3+{}^4c_4\mu_0'(\mu_1')^4$$

$$\mu_4=\mu_4'-4\mu_3'\mu_1'+6\mu_2'(\mu_1')^2-3(\mu_1')^4 \quad \left(\because {}^nc_n = 1 \text{ and } {}^nc_{n-1} = n\right)$$

$$\mu_4=\mu_4'-4\mu_1'\mu_3'+6(\mu_1')^2\mu_2'-3(\mu_1')^4 \qquad (\text{Eq. } 4.16)$$

4.2 Skewness:-

Literal meaning of skewness is lack of symmetry. It measures the degree of departure of a distribution from symmetry and reveals the direction of scatterdness of the items.

A frequency distribution is said to be symmetrical when values of the variables equidistant from their mean have equal frequencies. If a frequency distribution is not symmetrical, it is said to be asymmetrical or skewed. Any deviation from symmetry is called skewness.

According to **Morris Humberg** Skewness refers to the asymmetry or lack of symmetry in the shape of a frequency distribution.

According to **Croxton & Cowden** When a series is not symmetrical it is said to be asymmetrical or skewed.

According to **Simpson & Kafka** Measures of skewness tell us the direction and the extent of skewness. In a symmetrical distribution the mean, median and mode are identical. The more we move away from the mode, the larger the asymmetry or skewness.

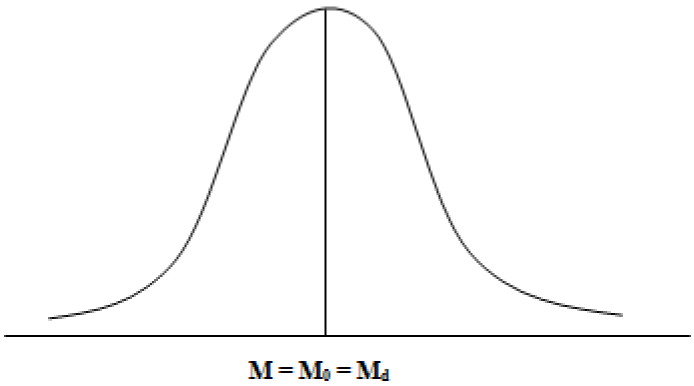
In the words of **Riggleman and Frisbee Skewness** is the lack of symmetry when a frequency distribution is plotted on a chart, skewness present in the items tends to be dispersed more on one side of the mean than on the other.

Thus the above definitions make it clear that the word skewness refers to the lack of symmetry. If a distribution is normal there would be no skewness in it and the curve drawn from the distribution would be symmetrical. In case of skewed distributions the curve drawn would be elongated either to the left or to the right. The concept of skewness gains importance from the fact that statistical theory is often based upon the assumption of the normal distribution. A measure of skewness is, therefore necessary in order to guard against the consequences of the assumption. The following three figures would give an idea about the shape of symmetrical and asymmetrical curves.

4.2.1 Symmetrical curve

The figure is given below, presents the shape of a symmetrical curve which is bell shaped having no skewness. The value of mean (M), median (M_d) and mode (M_o) for such a curve would be identical.

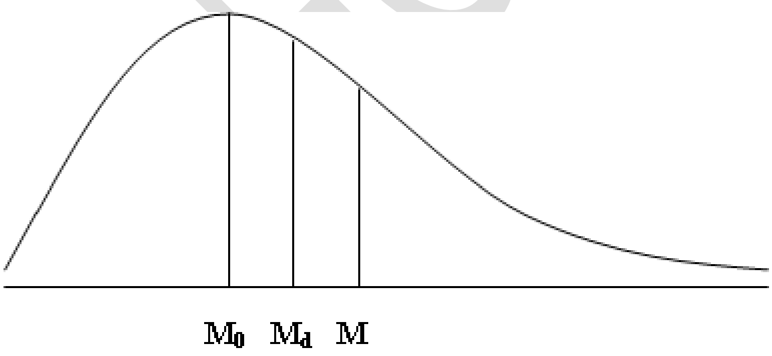
In a symmetrical distribution the values of mean, median and mode coincide. The spread of the frequencies is the same on both sides of the centre point of the curve. For a symmetrical distribution Mean = Median = Mode.



Symmetrical distribution

4.2.2 Positively skewed curve

A positively skewed curve has a longer tail towards the higher values of X i.e. the frequency curve gradually slopes down towards the higher values of X. In a positively skewed distribution the mean is greater than the median and then mode and the median lies in between mean and mode. The frequencies are spread over a greater range of values on the high value end of the curve (the right hand side). For a positively skewed distribution Mean > Median > Mode.

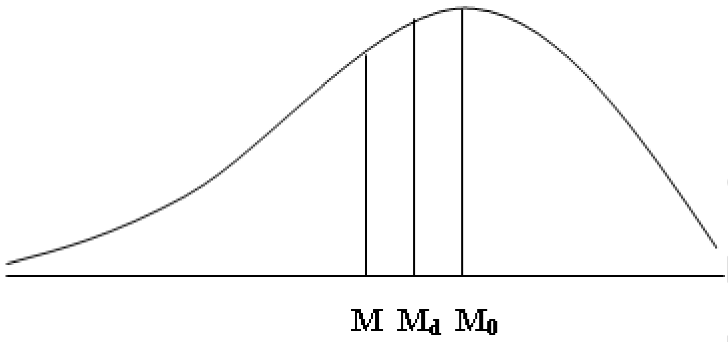


Positively skewed distribution

4.2.3 Negatively skewed curve

A negatively skewed curve has a longer tail towards the lower values of X i.e. the frequency curve gradually slopes down towards the lower values of X as shown in the following figure.

In the negatively skewed distribution the mode is the maximum and mean is the least. The median lies in between mean and mode. The elongated tail in negatively skewed distribution is on the left hand side as would be clear from the figure. For a negatively skewed distribution, Mean < Median < Mode.



Negatively skewed distribution

4.2.4 Measures of Skewness

Measures of skewness are meant to give an idea about the extent of asymmetry in a series. A distribution is said to be skewed if

[i] The frequency curve of the distribution is not a symmetric bell shaped curve but stretched more to one side than to the other.

[ii] The values of mean (M), median (M_d) and mode (M_o) fall at different points i.e they don't coincide.

[iii] Quartiles Q_1 and Q_3 are not equidistant from the median.

$Q_3 - Q_2 \neq Q_2 - Q_1$

[iv] The corresponding pairs of deciles and percentiles are not equidistant from the median.

If a particular distribution is found to be skewed, the next problem that arises is to measure the extent of skewness. To find out the direction and extent of asymmetry in a series statistical measures of skewness are employed.

These measures can be absolute or relative. The absolute measures of skewness tell us the extent of asymmetry and whether it is positive or negative. The absolute skewness is based on the difference between mean and mode. Symbolically,

Absolute skewness = Mean - Mode

Skewness will be positive, if the value of mean is greater than the mode and skewness will be negative, if the value of mean is less than the mode. The difference between the mean and the mode, whether positive or negative, indicates the distribution is positively skewed or negatively skewed. However, such an absolute measure of skewness is not adequate because it cannot be used for comparison of skewness in two distributions, if they are in different units, since difference between the mean and mode will be in terms of the units of distribution. Thus for comparison purposes we use the relative measures of skewness known as co-efficient of skewness.

4.2.5 Karl pearson's coefficient of skewness

The first coefficient of skewness is defined by Karl Pearson

$$\therefore \text{Karl Pearson's coefficient of skewness}\,(S_K) = \frac{\text{A.M.-Mode}}{\text{Standard deviation}} = \frac{\overline{X}\text{-Mo}}{\sigma}$$

This measure is based on the fact that the mean and the mode are drawn widely apart. Skewness will be positive if mean > mode and negative if mean < mode. **There is no limit to this measure in theory and this is a slight drawback.** But in practice the value given by this formula is rarely very high and its value usually lies between -1 and +1.

This coefficient is a pure number without units since both numerator and denominator have the same dimensions. The value of this coefficient lies between -3 and +3.

If mode is ill-defined. It cannot be computed, hence there is difficulty in computing Karl Pearson's coefficient of skewness. In such situation following empirical relation is used.

$$\left(\overline{X}\text{-mode}\right) \cong 3\left(\overline{X}\text{-median}\right)$$

$$\text{Mode} \cong \text{median-}2\overline{X}$$

4.2.6 Bowley's Coefficient of Skewness

Prof. A.L. Bowley's Coefficient of Skewness is also known as Coefficient of Skewness based on quartiles. This coefficient is a pure number without units since both numerator and denominator have the same dimensions. The value of this coefficient lies between -1 and +1. It is given by,

$$\text{Bowley's Coefficient of Skewness}(S_K) = \frac{(Q_3\text{-}Q_2)\text{-}(Q_2\text{-}Q_1)}{(Q_3\text{-}Q_2)+(Q_2\text{-}Q_1)} = \frac{Q_3\text{-}2Q_2+Q_1}{(Q_3\text{-}Q_1)}$$

This is especially useful in situations where quartiles and median are used viz.

[i] When the mode is ill-defined and extreme observations are present in the data.

[ii] When the distribution has open end classes or unequal class intervals.

4.2.7 Kelly's Coefficient of Skewness

The drawback of Bowley's Coefficient of Skewness is that it ignores the 50% of the data which can be partially removed by taking two deciles or percentiles equidistant from the median value. The refinement was suggested by Kelly.

$$\text{Kelly's Coefficient of Skewness}(S_K) = \frac{P_{90}+P_{10}\text{-2median}}{P_{90}\text{-}P_{10}} = \frac{D_9+D_1\text{-2median}}{D_9\text{-}D_1}$$

4.2.8 Pearsonian Coefficient of Skewness based on moments (β and γ coefficients)

Karl Pearson gave the following coefficients calculated from the moments about mean. These coefficients are pure numbers independent of units of measurement and as such can be conveniently used for comparative studies and it is defined as follows:

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} \qquad \text{and} \qquad \gamma_1 = \sqrt{\beta_1} = \frac{\mu_3}{\sqrt{\mu_2^3}}$$

Interpretation:

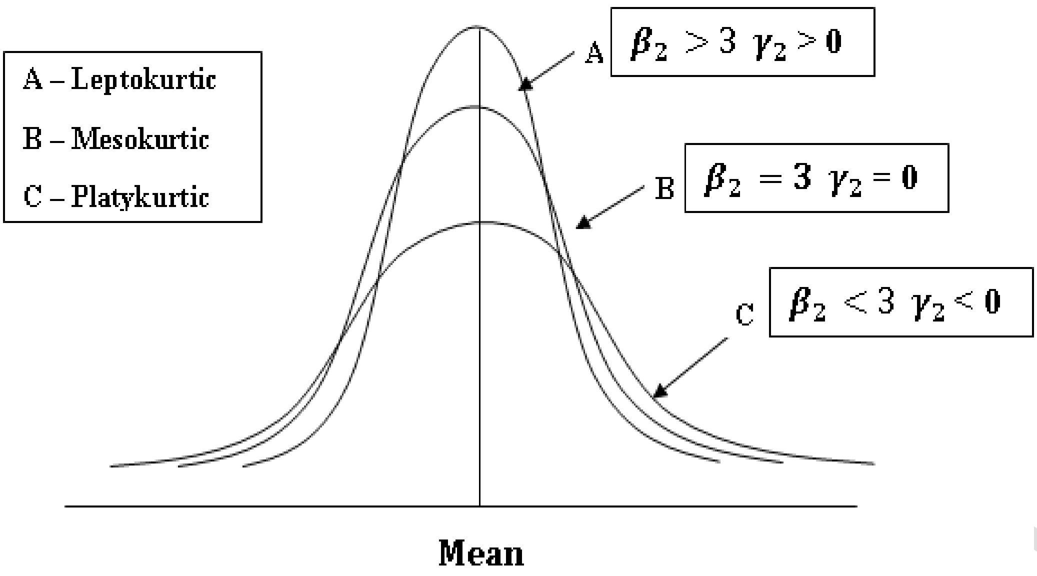
[1] If $\gamma_1 < 0$, the distribution is negatively skewed.

[2] If $\gamma_1 = 0$, the distribution is symmetric.

[3] If $\gamma_1 > 0$, the distribution is positively skewed.

4.3 Kurtosis:-

The expression Kurtosis is used to describe the peakedness of a curve. Kurtosis is a Greek word means bulginess. In statistics kurtosis refers to the degree of flatness or peakedness in the region about the mode of a frequency curve. The degree of Kurtosis of a distribution is measured relative to the peakedness of normal curve. If we know the measures of central tendency, dispersion and skewness, we cannot still form a complete idea about the distribution.



Type of kurtosis

All the three curves are symmetrical about mean and have same variation (range). In order to identify a distribution completely we need one more measure which Prof. Karl Pearson called convexity of the curve or kurtosis. It is measured by β_2 and γ_2 given as under

$$\beta_2 = \frac{\mu_4}{\mu_2^2} \text{ and } \gamma_2 = \beta_2 - 3 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{\mu_4 - 3\mu_2^2}{\mu_2^2}$$

Curve of type B which is neither flat nor peaked is known as normal curve and shape of its hump is accepted as a standard one. Curves with humps of the form of normal curve are said to have normal kurtosis and are termed as **Mesokurtic** $\beta_2 = 3$ and $\gamma_2 = \beta_2 - 3 = 0$ .The curve of type A , which are more peaked than the normal curve are known as **Leptokurtic** $\beta_2 > 3$ and $\gamma_2 > 0$ and are said to lack kurtosis. Curves of type C, which are flatter than the normal curve are called **Platykurtic** $\beta_2 < 3$ and $\gamma_2 < 0$ and they are said to posses kurtosis in excess.

Note:

[i] γ_2 ,is called kurtosis or excess of kurtosis.

[ii] β_2 and γ_2, are invariant to change of origin and scale.

[iii] β_2 and γ_2, are both free from units.

[iii] β_2 and γ_2, **cannot be used for qualitative data and frequency distribution having open end classes. In such situation quartiles and percentiles are used. It is denoted by**

$$K_u = \frac{(Q_3 - Q_1)/2}{(P_{90} - P_{10})}$$

For normal distribution $K_u = 0.263$

Properties of Central Moments:-

[1] **Effect of change of origin: The central moments are invariant to the change of origin. In other words, If** $u = x - a$ then μ_r of $u = \mu_r$ of x

Proof : Let $u = x - a$, hence $\bar{u} = \bar{x} - a$

By defination

$$\mu_r \text{ of } u = \frac{\sum (u_i - \bar{u})^r}{n}$$

$$\mu_r \text{ of } u = \frac{\sum \left[(x_i - a) - (\bar{x} - a)\right]^r}{n}$$

$$\mu_r \text{ of } u = \frac{\sum \left[(x_i - \bar{x})\right]^r}{n}$$

$$\mu_r \text{ of } u = \mu_r \text{ of } x$$

[2] Effect of change of origin and scale: $u = \frac{x-a}{h}$ then μ_r of $u = \frac{1}{h^r}\mu_r$ of x

Proof: Since, $u = \frac{x-a}{h}$, we get $\bar{u} = \frac{\bar{x}-a}{h}$

By defination,

$$\mu_r \text{ of } u = \frac{\sum (u_i - \bar{u})^r}{n}$$

$$\mu_r \text{ of } u = \frac{1}{n}\sum\left[\frac{x_i - a}{h} - \frac{\bar{x} - a}{h}\right]^r$$

$$\mu_r \text{ of } u = \frac{1}{h^r}\frac{\sum(x_i - \bar{x})^r}{n}$$

$$\mu_r \text{ of } u = \frac{1}{h^r}\mu_r \text{ of } x$$

Result on Skewness:

[1] The Pearsonian coefficient $\beta_2 \geq 1$

Proof: Suppose $x_1, x_2, \ldots, x_n$ are the observations then second and fourth central moments are

$$\mu_2 = \frac{\sum(x_i - \bar{x})^2}{n} \qquad \text{and} \qquad \mu_4 = \frac{\sum(x_i - \bar{x})^4}{n}$$

$$\text{Let } (x_i - \bar{x})^2 = y_i$$

Note that,

$$\mathrm{Var}(y) \geq 0$$

$$\therefore \frac{\sum y_i^2}{n} - (\bar{y})^2 \geq 0$$

$$\frac{1}{n}\sum y^2 - \left(\frac{1}{n}\sum y\right)^2 \geq 0$$

$$\therefore \frac{\sum(x_i - \bar{x})^2}{n} - \left[\frac{\sum(x_i - \bar{x})^2}{n}\right]^2 \geq 0$$

$$\mu_4 - \mu_2^2 \geq 0$$

$$\therefore \mu_4 \geq \mu_2^2$$

$$\therefore \frac{\mu_4}{\mu_2^2} \geq 1;$$

$$\therefore \beta_2 \geq 1$$

[2] If β_1 and β_2 **are the Pearsonian coefficient then** $\beta_2 \geq \beta_2 + 1$

Proof: Suppose $x_1, x_2, \ldots, x_n$ are the observations

$$y_i = x_i - \bar{x}, \text{ hence, } \mu_r = \frac{\sum y_i^r}{n}$$

Note that,

$$\frac{1}{n}\sum(ay_i^{\,2}+by_i+c)^2 \geq 0$$

$$\therefore \frac{1}{n}\left(a^2\sum y_i^{\,4}+b^2\sum y_i^{\,2}+\sum c^2+2ab\sum y_i^{\,3}+2bc\sum y_i+2ac\sum y_i^{\,2}\right)\geq 0$$

$$\therefore \mu_4a^2+\mu_2b^2+c^2+2ab\mu_3+2bc\mu_1+2ac\mu_2 \geq 0 \qquad ....(1)$$

We use the following result:The expression

$$Ax^2+by^2+cz^2+2Hxy+2Gxz+2Fyz \geq 0 \quad ....(2)$$

$$\begin{vmatrix} A & H & G \\ H & B & F \\ G & F & C \end{vmatrix} \geq 0$$

Comparing (1) and (2), we get

$A=\mu_4, B=\mu_2, C=1, H=\mu_3, G=\mu_2, F=\mu_1=0$

Expression is always non-negative if

$$\begin{vmatrix} \mu_4 & \mu_3 & \mu_2 \\ \mu_3 & \mu_2 & 0 \\ \mu_2 & 0 & 1 \end{vmatrix} \geq 0$$

$$\therefore \mu_4\mu_2 - \mu_2^3+\mu_2(-\mu_2^2) \geq 0$$

$$\therefore \mu_4\mu_2 \geq \mu_3^2+\mu_2^3 \qquad \text{(dividing by } \mu_2^3\text{)}$$

$$\therefore \frac{\mu_4}{\mu_2^2} \geq \frac{\mu_3^2}{\mu_2^3}+1$$

$$\therefore \beta_2 \geq \beta_1+1 \text{ or } \beta_2-\beta_1-1 \geq 0$$

Problem on Moments, Skewness and Kurtosis

[1] a) The standard deviation of symmetrical distribution is 3 . What must be the value of the fourth moment about the mean in order that the distribution be mesokurtic?

b) If the first four moments of distribution about the value 5 are equal to -4 , 22 ,-117 and 560 determine the corresponding moments about the mean

Solution:-

(a) For a mesokurtic distribution

We are given

$\sigma = 3$ and $\beta_2 = 3$

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

$\beta_2 = \frac{\mu_4}{{\mu_2}^2}$ $\quad \left(\because \sigma^2 = \mu_2 = (3)^2 = 9, \right)$

$3 = \frac{\mu_4}{9^2}$ $\quad \left(\because \beta_2 = 3 \right)$

$\therefore \mu_4 = 243$

Thus the fourth moment about mean must be 243 in order that distribution be mesokurtic.

b) We are given moments about an arbitrary origin 5. Thus,

$\mu_1' = -4,\ \mu_2' = 22,\ \mu_3' = -117,\ \mu_4' = 560$

Moments about mean:-

From these we can find out moments about mean from the following relationships:

$\mu_2 = \mu_2' - (\mu_1')^2 \qquad \Rightarrow \mu_2 = 22-(-4)^2$

$\mu_2 = 22-16 \qquad \Rightarrow \mu_2 = 6$

$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1')^3 \qquad \Rightarrow \mu_3 = -117-3(-4)(22)+2(-4)^3$

$\mu_3 = -117+264-128 \qquad \Rightarrow \mu_3 = 19$

$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4 \qquad \Rightarrow \mu_4 = 560-4(-4)(117)+6(22)(-4)^2-3(-4)^4$

$\mu_4 = 560-1872+2112-768 \qquad \Rightarrow \mu_4 = 32$

Thus the moments about mean are

$\mu_1 = 0,\ \mu_2 = 6,\ \mu_3 = 19$ and $\mu_4 = 32$

[2] You are given the following value of moment;

$\mu_2 = 44.553$; $\mu_3 = -9.774$; $\mu_4 = 5508.567$

Find the corrected values of each one of these taking in to account the class interval which is 3.

Solution:-

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

$$\mu_2(\text{corrected}) = 43.353 - \frac{(3)^2}{12}$$

$$\mu_2(\text{corrected}) = 43.353 - 0.75$$

$$\mu_2(\text{corrected}) = 42.603$$

$$\mu_4(\text{corrected}) = \mu_4(\text{uncorrected}) - \frac{1}{12} i^2 \mu_2(\text{uncorrected}) + \frac{7i^4}{240}$$

$$\mu_4(\text{corrected}) = 5508.567 - \frac{1}{12}(3)^2(43.353) + \frac{7(3)^4}{240}$$

$$\mu_4(\text{corrected}) = 5508.567 - 32.515 + 2.3625$$

$$\mu_4(\text{corrected}) = 5478.41$$

$\Rightarrow$ corrected values of $\mu_2 = 42.603$ and $\mu_4 = 5478.41$

Skewness

[3] Find Bowley's coefficient of skewness for the following frequency distribution:

No. of children per family	0	1	2	3	4	5	6
No. of families	7	10	16	25	18	11	8

Solution:-

No. of children per family (X)	No. of families (F)	c .f.

0	7	7
1	10	17
2	16	33
3	25	58
4	18	76
5	11	87
6	8	95

$$S_{KB} = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

$$Q_i = \text{Size of}\left(\frac{i(N+1)}{4}\right)^{th} \text{item of the series, } i = 1, 2, 3$$

$$Q_1 = \text{Size of}\left(\frac{1(N+1)}{4}\right)^{th} \text{item of the series}$$

$$Q_1 = \text{Size of}\left(\frac{1(95+1)}{4}\right)^{th}$$

$$Q_1 = \text{Size of}\left(\frac{96}{4}\right)^{th}$$

$$Q_1 = 24^{th} \text{item of the series}$$

See the cumulative frequency which is equal to Q_1 or just greater than Q_1 is 33. Corresponding to 33 variable value is 2 and therefore, $Q_1 = 2$

Now for Q_2

$$Q_2 = \text{Size of}\left(\frac{2(N+1)}{4}\right)^{th} \text{item of the series}$$

$$Q_2 = \text{Size of}\left(\frac{2(95+1)}{4}\right)^{th}$$

$$Q_2 = \text{Size of}\left(\frac{192}{4}\right)^{th}$$

$$Q_2 = 48^{th} \text{item of the series}$$

See the cumulative frequency which is equal to Q_2 or just greater than Q_2 is 58. Corresponding to it variable value is 3 and therefore, $Q_2 = 3$

Now for Q_3

$$Q_3 = \text{Size of} \left(\frac{3(N+1)}{4}\right)^{th} \text{item of the series}$$

$$Q_3 = \text{Size of} \left(\frac{3(95+1)}{4}\right)^{th}$$

$$Q_3 = \text{Size of} \left(\frac{288}{4}\right)^{th}$$

$$Q_3 = 72^{th} \text{item of the series}$$

See the cumulative frequency which is equal to Q_3 or just greater than Q_3 is 76. Corresponding to 76 variable value is 4 and therefore, $Q_3 = 4$

$$S_{KB} = \frac{4+2-2\times 3}{4-3}$$

$$S_{KB} = \frac{6-6}{1}$$

$$S_{KB} = 0$$

Bowley's coefficient of skewness = 0

[4] Calculate Bowley's coefficient of skewness for the following data

Annual sales (Rs'000)	No. of firms
Less than 20	30
Less than 30	225
Less than 40	165
Less than 50	580
Less than 60	634
Less than 70	644
Less than 80	650
Less than 90	665
Less than 100	680

Solution:-

Sales	F	Cum. Freq. (F)

10-20	30	30
20-30	195	225
30-40	240	465
40-50	115	580
50-60	54	634
60-70	10	644
70-80	6	650
80-90	15	605
90-100	15	680

For quartile first

$$Q_1 = \text{size of } \frac{N}{4}$$

$$Q_1 = \text{size of } \frac{680}{4}$$

$$Q_1 = \text{size of } 170$$

Cumulative frequency which is equal to 170 or just greater than 170 is 225.

First quartile class corresponding to 225 is 20-30. Hence,

$$Q_1 = l + \frac{h}{f_k}\left(\frac{N}{4} - F_{k-1}\right)$$

Where, l = lower limit of the median class = 20

f_k = frequency corresponding to the median class = 195

F_{k-1} = cumulative frequency of the class preceding to the first quartile class is 30 and h = 10

$$Q_1 = l + \frac{h}{f_k}\left(\frac{N}{4} - F_{k-1}\right)$$

$$Q_1 = 20 + \frac{10}{195}(170 - 30)$$

$$Q_1 = 20 + \frac{10}{195}(140)$$

$$Q_1 = 20 + \frac{1400}{195}$$

$$Q_1 = 20 + 7.18$$

$$Q_1 = 27.18$$

For quartile second

$$\text{Median} = \text{size of } \frac{2N}{4}$$

$$\text{Median} = \text{size of } \frac{680}{2}$$

$$\text{Median} = \text{size of } 340$$

Cumulative frequency which is, just greater than 340 is 465.

Median class corresponding to 465 is 30-40. Hence,

$$\text{Median} = l + \frac{h}{f_k}\left(\frac{2N}{4} - F_{k-1}\right)$$

Where,

l = lower limit of the median class = 30

f_k = frequency corresponding to the median class = 240

F_{k-1} = cumulative frequency of the class preceding to the median class=225

and h = 10

$$\text{Median} = l+\frac{h}{f_k}\left(\frac{2N}{4}-F_{k-1}\right)$$

$$\text{Median} = 30+\frac{10}{240}(340 - 225)$$

$$\text{Median} = 30+\frac{1}{24}(115)$$

$$\text{Median} = 30+4.79$$

$$\text{Median} = 34.79$$

For quartile third

$$Q_3 = \text{size of } \frac{3N}{4}$$

$$Q_3 = \text{size of } \frac{3\times 680}{4}$$

$$Q_3 = \text{size of } 510$$

Cumulative frequency which is equal to 210 or just greater than 510 is 580.

Third quartile class corresponding to 580 is 40-50. Hence,

$$Q_3 = l+\frac{h}{f_k}\left(\frac{3N}{4}-F_{k-1}\right)$$

Where,

l = lower limit of the median class = 40

f_k = frequency corresponding to the median class = 115

F_{k-1} = cumulative frequency of the class preceding to the first quartile class is 465 and h = 10

$$Q_1 = l + \frac{h}{f_k}\left(\frac{3N}{4} - F_{k-1}\right)$$

$$Q_3 = 40 + \frac{10}{115}(510 - 465)$$

$$Q_3 = 40 + \frac{10}{115}(45)$$

$$Q_3 = 40 + \frac{450}{115}$$

$$Q_3 = 40 + 3.91$$

$$Q_3 = 43.91$$

$$\text{Bowley's Coefficient of Skewness} = \frac{43.9+27.18-2(34.79)}{43.9-27.18}$$

$$\text{Bowley's Coefficient of Skewness} = \frac{71.08 - 69.58}{16.72}$$

$$\text{Bowley's Coefficient of Skewness} = \frac{1.5}{16.72}$$

$$\text{Bowley's Coefficient of Skewness} = 0.09$$

[5] In a frequency distribution the coefficient of skewness based on quartiles is 0.6. If the sum of the upper and the lower quartiles is 100 and the median is 38, find the value of the upper quartile.

Solution: we are given that

Coeff. of $S_k = 0.6$, $Q_1 + Q_3 = 100$

Median = Q_2 = 38

We have to find Q_3

$$S_{KB} = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

$$0.6 = \frac{100 - 2\times 38}{Q_3 - Q_1}$$

$$0.6\,(Q_3 - Q_1) = 100 - 76 \qquad \ldots(i)$$

Since, $Q_3 + Q_1 = 100 \Rightarrow Q_1 = 100 - Q_3$

Putting value of Q_1 in equation (i),we get

$0.6\left[Q_3 - (100 - Q_3)\right] = 24$

$\left[2Q_3 - 100\right] = 40$

$2Q_3 = 140$

$Q_3 = 70$

$\therefore$ Upper quartile is 70

[6] Given Q_1=18,Q_3 = 25, Mode=21, Mean =18. Find the coefficient of skewness.

Solution:

$$S_{KB} = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$$

Mode = 3Med-2 $\bar{x}$

21=3Med-2 x 18

3 Med= 21+36

3 Med= 57

Median =19

$$S_{KB} = \frac{25+18-2\times 19}{25-18}$$

$$S_{KB} = \frac{43-38}{7}$$

$$S_{KB} = \frac{5}{7}$$

S_{KB} =0.714

$\therefore$ Coefficient of skewness = 0.714

[7] For a moderately skewed data. The arithmetic mean is 200. The coefficient of variation is 8 and Karl Pearson's coefficient of skewness is 0.3 find the mode and median.

Solution:- We are given that

$\bar{x}$ = 200, C .V. = 8, S_{kp} = 0.3

We have to find mode and median from the formula of coefficient variation, we can determine the S.D.

$$\text{Coefficient of Variation} = \frac{\sigma}{\bar{x}} \times 100$$

$$8 = \frac{\sigma}{200} \times 100$$

$$\sigma = \frac{8}{100} \times 200$$

$$\sigma = 16$$

$$\text{Coefficient of } S_K = \frac{\text{Mean-mode}}{\sigma}$$

$$0.3 = \frac{200\text{-mode}}{16}$$

200-mode = 4.8

200-4.8 = Mode

Mode = 195.2

and

Mode = 3 Med – 2 Mean

195.2 = 3 Med – 2 (200)

3 Med – 400 = 195.2

3 Med = 595.2

Median = 198.4

Mode = 195.2 and Median = 198.4

[8] Calculate the coefficient of skewness from the following data

Mid- Point	15	20	25	30	35	40
Frequency	12	18	25	24	20	21

Solution:-

Mid- Point (m)	Frequency (f)	$d_i = \frac{m_i - A}{h}$ $d_i = \frac{m_i - 25}{5}$	$f_i d_i$	$f_i d_i^2$

15	12	-2	-24	45
20	18	-1	-18	18
25	25	0	0	0
30	24	1	24	24
35	20	2	40	80
40	21	3	63	189
Total	120		85	359

$$\text{Coefficient of } S_K = \frac{\text{Mean-mode}}{\sigma}$$

Now for Mean,

$$\bar{x} = A + \frac{h\sum f_i d_i}{N}$$

$$\bar{x} = 25 + \frac{5\times 85}{120}$$

$$\bar{x} = 25 + \frac{425}{120}$$

$$\bar{x} = 25 + 3.54$$

$$\bar{x} = 28.54$$

Now for Standard deviation

$$\sigma = \sqrt{\frac{\sum f_i d_i^{\,2}}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h$$

$$\sigma = \sqrt{\frac{359}{120} - \left(\frac{85}{120}\right)^2} \times 5$$

$$\sigma = \sqrt{2.99 - (0.70)^2} \times 5$$

$$\sigma = \sqrt{2.99 - 0.49} \times 5$$

$$\sigma = \sqrt{2.5} \times 5$$

$$\sigma = 1.58 \times 5$$

$$\sigma = 7.9$$

Now for Mode,

It can be clearly seen that mid-point 25 has occurred maximum number of times means, it has maximum frequency i. e. 25

Hence, corresponding to 25, class is 22.5-27.5, it is called modal class

Where,

l = lower limit of the modal class = 22.5

f_1 = frequency of the modal class = 25

f_0 =frequency of the class preceding to the modal class = 18

f_2 = frequency of the class succeeding to the modal class = 24

h = size of the class = 5

$$\text{Mode }(M_o) = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$$

$$\text{Mode }(M_o) = 22.5 + \frac{25-18}{2 \times 25-18-24} \times 5$$

$$\text{Mode }(M_o) = 22.5 + \frac{7}{50-42} \times 5$$

$$\text{Mode }(M_o) = 22.5 + \frac{35}{8}$$

$$\text{Mode }(M_o) = 22.5 + 4.37$$

$$\text{Mode }(M_o) = 26.87$$

Now,

$$\text{Coefficient of } S_K = \frac{\text{Mean-mode}}{\sigma}$$

$$\text{Coefficient of } S_K = \frac{28.54 - 26.87}{7.9}$$

$$\text{Coefficient of } S_K = \frac{1.67}{7.9}$$

$$\text{Coefficient of } S_K = 2.11$$

[9] From the following data of age of employees, calculate coefficient of skewness and comment on the result:

Age below	25	30	35	40	45	50	55
No. of employees	8	20	40	65	80	92	100

Solution:

Age	Mid-Point (m)	Frequency (f)	$d_i = \frac{m_i - A}{h}$ $d_i = \frac{m_i - 37.5}{5}$	$f_i d_i$	$f_i d_i^2$
20-25	22.5	8	-3	-24	72
25-30	27.5	12	-2	-24	48
30-35	32.5	20	-1	-20	20
35-40	37.5	25	0	0	0
40-45	42.5	15	1	15	15
45-50	47.5	12	2	24	48
50-55	52.5	8	3	24	72
		100		-5	

$$\text{Coefficient of } S_K = \frac{\text{Mean-mode}}{\sigma}$$

Now for Mean,

$$\bar{x} = A + \frac{h\sum f_i d_i}{N}$$

$$\bar{x} = 37.5 + \frac{5\times(-5)}{100}$$

$$\bar{x} = 37.5 - \frac{25}{100}$$

$$\bar{x} = 37.5 - 0.25$$

$$\bar{x} = 37.25$$

Mode: By inspection mode lies in the class 35-40

$$\text{Mode }(M_o) = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$$

$$\text{Mode }(M_o) = 35 + \frac{25-20}{2\times 25-20-15} \times 5$$

$$\text{Mode }(M_o) = 35 + \frac{25}{15}$$

$$\text{Mode }(M_o) = 35 + 1.67$$

$$\text{Mode }(M_o) = 36.67$$

Now for Standard deviation

$$\sigma = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h$$

$$\sigma = \sqrt{\frac{275}{100} - \left(\frac{5}{100}\right)^2} \times 5$$

$$\sigma = \sqrt{2.75 - (0.05)^2} \times 5$$

$$\sigma = \sqrt{2.75 - 0.0025} \times 5$$

$$\sigma = \sqrt{2.7475} \times 5$$

$$\sigma = 1.66 \times 5$$

$$\sigma = 8.3$$

Now,

$$\text{Coefficient of } S_K = \frac{\text{Mean-mode}}{\sigma}$$

$$\text{Coefficient of } S_K = \frac{37.25 - 36.67}{8.3}$$

$$\text{Coefficient of } S_K = \frac{0.58}{8.3}$$

$$\text{Coefficient of } S_K = 0.070$$

$\therefore$ **Karl Pearson's Coefficient of Skewness is 0.07**

[10] Find Karl Pearson's coefficient of skewness for the following distribution and comment on the result:

Class	3-7	8-12	13-17	18-22	23-27	28-32

Frequency	108	580	175	80	32	18

Solution:

Class	Mid- Point (m)	Frequency (f)	$d_i = \frac{m_i\text{-}A}{h}$ $d_i = \frac{m_i\text{-}15}{5}$	$f_i d_i$	$f_i d_i^2$	c.f.
2.5-7.5	5	108	-2	216	432	108
7.5-12.5	10	580	-1	-580	580	688
12.5-17.5	15	175	0	0	0	863
17.5-22.5	20	80	1	80	80	943
22.5-27.5	25	32	2	64	128	975
27.5-32.5	30	18	3	54	162	993

$$\text{Coefficient of } S_K = \frac{\text{Mean-mode}}{\sigma}$$

Now for Mean,

$$\bar{x} = A + \frac{h\sum f_i d_i}{N}$$

$$\bar{x} = 15 + \frac{5\times(-598)}{993}$$

$$\bar{x} = 15 - \frac{2990}{993}$$

$\bar{x}$ = 15 -3.01

$\bar{x}$ = 11.99

Mean =11.99

Mode: By inspection mode lies in the class $7.5 - 12.5$

$$\text{Mode }(M_o) = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$$

$$\text{Mode }(M_o) = 7.5 + \frac{580-108}{2 \times 580-108-175} \times 5$$

$$\text{Mode }(M_o) = 7.5 + \frac{472}{1160-283} \times 5$$

$$\text{Mode }(M_o) = 7.5 + \frac{2360}{877}$$

$$\text{Mode }(M_o) = 7.5 + 2.69$$

$$\text{Mode }(M_o) = 10.19$$

Mode= 10.19

Now for Standard deviation

$$\sigma = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h$$

$$\sigma = \sqrt{\frac{1382}{993} - \left(\frac{-598}{993}\right)^2} \times 5$$

$$\sigma = \sqrt{1.39 - (-0.60)^2} \times 5$$

$$\sigma = \sqrt{1.39 - 0.36} \times 5$$

$$\sigma = \sqrt{1.03} \times 5$$

$$\sigma = 1.01 \times 5$$

$$\sigma = 5.05$$

Now,

$$\text{Coefficient of } S_K = \frac{\text{Mean-mode}}{\sigma}$$

$$\text{Coefficient of } S_K = \frac{11.99 - 10.19}{5.05}$$

$$\text{Coefficient of } S_K = \frac{1.80}{5.05}$$

$$\text{Coefficient of } S_K = 0.36$$

Karl Pearson's coefficient of skewness is 0.36

[11] Calculate Karl Pearson's coefficient of skewness from the following data

Marks more than	5	15	25	35	45	55	65	75	85
No. of student	120	105	96	85	72	58	42	12	0

Solution:-

Marks	Mid-Point (m)	Frequency (f)	$d_i = \frac{m_i - A}{h}$ $d_i = \frac{m_i - 50}{10}$	$f_i d_i$	$f_i d_i^2$	c.f.
5-15	10	15	-4	-60	240	15
15-25	20	9	-3	-27	81	24
25-35	30	11	-2	-22	44	35
35-45	40	13	-1	-13	13	48
45-55	50	14	0	0	0	62
55-65	60	16	1	16	16	78
65-75	70	30	2	60	120	108
75-85	80	12	3	36	108	120

$$\text{Coefficient of } S_K = \frac{3(\text{Mean-Median})}{\sigma}$$

Now for Mean,

$$\bar{x} = A + \frac{h\sum f_i d_i}{N}$$

$$\bar{x} = 50 + \frac{10\times(-10)}{120}$$

$$\bar{x} = 50 - \frac{100}{120}$$

$$\bar{x} = 50 - 0.83$$

$$\bar{x} = 49.17$$

Mean = 49.17

For median

$$\text{Median} = \text{size of } \frac{2N}{4}$$

$$\text{Median} = \text{size of } \frac{240}{4}$$

$$\text{Median} = \text{size of } 60$$

Cumulative frequency which is equal to 60 or just greater than 60 is 62.

Median class corresponding to 62 is 45-55. Hence,

$$\text{Median} = l + \frac{h}{f_k}\left(\frac{2N}{4} - F_{k-1}\right)$$

Where,

l = lower limit of the median class = 45

f_k = frequency corresponding to the median class = 14

F_{k-1} = cumulative frequency of the class preceding to the median class= 48 and h = 10

$$\text{Median} = l + \frac{h}{f_k}\left(\frac{2N}{4} - F_{k-1}\right)$$

$$\text{Median} = 45 + \frac{10}{14}(60 - 48)$$

$$\text{Median} = 45 + \frac{120}{14}$$

$$\text{Median} = 45 + 8.57$$

$$\text{Median} = 53.57$$

Now for Standard deviation

$$\sigma = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h$$

$$\sigma = \sqrt{\frac{622}{120} - \left(\frac{-10}{120}\right)^2} \times 10$$

$$\sigma = \sqrt{5.18 - (-0.083)^2} \times 10$$

$$\sigma = \sqrt{5.18 - 0.006889} \times 10$$

$$\sigma = \sqrt{5.17} \times 10$$

$$\sigma = 2.275 \times 10$$

$$\sigma = 22.75$$

$$\text{Coefficient of } S_K = \frac{3(\text{Mean-Median})}{\sigma}$$

$$\text{Coefficient of } S_K = \frac{3(49.17 - 53.57)}{22.75}$$

$$\text{Coefficient of } S_K = -\frac{13.2}{22.75}$$

$$\text{Coefficient of } S_K = -0.58$$

$\therefore$ Karl Pearson’s coefficient of skewness is -0.58

[12] Calculate coefficient of skewness by Karl Pearson`s method and the value of β_1 and β_2 form the following data:

Protits(Rs.laks)	10-20	20-30	30-40	40-50	50-60
No of companies	18	20	30	22	10

Solution:

Profits Rs. lakhs	Mid-Point (m)	Frequency (f)	$d_i = \frac{m_i - A}{h}$ $d_i = \frac{m_i - 35}{10}$	$f_i d_i$	$f_i d_i^2$	$f_i d_i^3$	$f_i d_i^4$

10-20	15	18	-2	-36	72	-144	288
20-30	25	20	-1	-20	20	-20	20
30-40	35	30	0	0	0	0	0
40-50	45	22	1	22	22	22	22
50-60	55	10	2	20	40	80	160
Total		100		-14	154	-62	490

Coefficient of $S_K = \frac{\text{Mean-mode}}{\sigma}$

Now for Mean,

$$\bar{x} = A + \frac{h\sum f_i d_i}{N}$$

$$\bar{x} = 35 - \frac{10 \times 14}{100}$$

$$\bar{x} = 35 - \frac{140}{100}$$

$$\bar{x} = 35 - 1.4$$

$$\bar{x} = 33.6$$

Now for Standard deviation

$$\sigma = \sqrt{\frac{\sum f_i d_i^2}{N} - \left(\frac{\sum f_i d_i}{N}\right)^2} \times h$$

$$\sigma = \sqrt{\frac{154}{100} - \left(\frac{-14}{100}\right)^2} \times 10$$

$$\sigma = \sqrt{1.54 - (-0.14)^2} \times 10$$

$$\sigma = \sqrt{1.54 - 0.0196} \times 10$$

$$\sigma = \sqrt{1.5204} \times 10$$

$$\sigma = 1.233 \times 10$$

$$\sigma = 12.33$$

Now for Mode,

It can be clearly seen that class 30-40 has occurred maximum number of times means, it has maximum frequency i. e. 30. Hence, corresponding to 30, class is 30-40, it is called modal class

Where,

l = lower limit of the modal class = 30

f_1 = frequency of the modal class = 30

f_0 =frequency of the class preceding to the modal class = 20

f_2 = frequency of the class succeeding to the modal class = 22

h = size of the class = 10

$$\text{Mode } (M_o) = l + \frac{f_1 - f_0}{2f_1 - f_0 - f_2} \times h$$

$$\text{Mode } (M_o) = 30 + \frac{30-20}{2 \times 30-20-22} \times 10$$

$$\text{Mode } (M_o) = 30 + \frac{10}{60-42} \times 10$$

$$\text{Mode } (M_o) = 30 + \frac{100}{18}$$

$$\text{Mode } (M_o) = 30 + 5.55$$

$$\text{Mode } (M_o) = 35.55$$

Now,

$$\text{Coefficient of } S_K = \frac{\text{Mean-Mode}}{\sigma}$$

$$\text{Coefficient of } S_K = \frac{33.6 - 35.56}{12.33}$$

$$\text{Coefficient of } S_K = \frac{-1.96}{12.33}$$

$$\text{Coefficient of } S_K = -0.159$$

Calculate β_1 and β_2

$$\mu_1^{'}=\frac{\sum fd}{N}$$

$$\mu_1^{'}=\frac{-14}{100}$$

$$\mu_1^{'}=-0.14$$

$$\mu_2^{'}=\frac{\sum fd^2}{N}$$

$$\mu_2^{'}=\frac{154}{100}$$

$$\mu_2^{'}=1.54$$

$$\mu_3^{'}=\frac{\sum fd^3}{N}$$

$$\mu_3^{'}=\frac{-62}{100}$$

$$\mu_3^{'}=-0.62$$

$$\mu_4^{'}=\frac{\sum fd^4}{N}$$

$$\mu_4^{'}=\frac{490}{100}$$

$$\mu_4^{'}=4.9$$

$\mu_2=\mu'_2-(\mu'_1)^2$

$\mu_2=1.54-(-0..14)^2$

$\mu_2=1.54-0.0196$

$\mu_2=1.5204$

$\mu_2=1.5204$

$\mu_3=\mu'_3-3\mu'_2\mu'_1+2(\mu'_1)^3$

$\mu_3=-0.62+0.6468-0.0055$

$\mu_3=0.0213$

$\mu_3=0.0213$

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

$$= 4.9 - 4(-0.14)(-0.62) + 6(-0.14)^2(1.54) - 3(-0.14)^4$$

$$= 4.9 - 0.3472 + 0.1811 + 0.0012$$

$$= 4.735$$

$$\mu_4 = 4.735$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\sqrt{\mu_2^{\ 3}}} \qquad \Rightarrow \sqrt{\beta_1} = \frac{0.0213}{\sqrt{(1.5204)^3}}$$

$$\sqrt{\beta_1} = \frac{0.0213}{\sqrt{3.52}} \qquad \Rightarrow \beta_1 = \frac{0.0213}{1.8747}$$

$$\beta_1 = 0.0114$$

$$\beta_2 = \frac{4.735}{(1.5204)^2} \qquad \Rightarrow \beta_2 = \frac{4.735}{2.312}$$

$$\beta_2 = 2.048$$

Karl Pearson's Coefficient of skewness $= -0.159$

$\beta_1 = 0.0114$ and $\beta_2 = 2.048$

[13] In a certain distribution the following result were obtained: $\bar{x} = 45$, median = 48, coefficient of skewness = -0.4. Find the value of s.d. and you are required to estimate it with the help of the available information.

Solution:- we have given that

coefficient of skewness =-0.4;   Median=48; Mean ($\bar{x}$) = 45

$$\text{Coefficient of skewness} = \frac{3(\text{mean-median})}{\sigma}$$

$$-0.4 = \frac{3(45-48)}{\sigma}$$

$$-0.4 = \frac{3(-3)}{\sigma}$$

$$-0.4 = \frac{-9}{\sigma}$$

$$\sigma = \frac{-9}{-0.4}$$

$\sigma = 22.5$

$\therefore$ standard deviation = $\sigma = 22.5$

[14] From the information given below calculate Karl Pearson's coefficient of Skewness.

Measure	Place A	Place B
Mean	256.5	240.0
Median	201.0	201.6
S.D.	215.0	181.0

Solution:-

We have to determine Karl Pearson's coefficient of Skewness.

$$\text{Coefficient of } S_K = \frac{\text{Mean-mode}}{\sigma}$$

We have given the information about mean, median, standard deviation of Place A and Place B

First find mode for Place A, median=201.1, Mean=256.5,

S.D.= 215.0

Mode = 3(Median)-2(Mean)

Mode = 3(201.1)-2(256.5)

= 603.3-513

Mode = 90.3

$$\text{Coefficient of } S_K = \frac{256.5\text{-}90.3}{215}$$

$$\text{Coefficient of } S_K = \frac{166.2}{215}$$

$\text{Coefficient of } S_K = 0.77$

Find mode for Place B, Median=201.6,Mean=240.8, S.D.=215.0

Mode = 3(Median)-2(Mean)

Mode=3(201.6)-2(240.8)

$=604.8-481.6$

Mode=123.2

$$\text{Coefficient of } S_K = \frac{240.8-123.2}{181}$$

$$\text{Coefficient of } S_K = \frac{117.6}{181}$$

Coefficient of $S_K = 0.65$

Karl Pearson's coefficient of Skewness for Place A is 0.774

Karl Pearson's coefficient of Skewness for Place B is 0.65

[15] Calculate first four moments about the mean and also the value of β_1 and β_2 from the following data

Marks	0-10	10-20	20-30	30-40	40-50	50-60	60-70
No. of marks	8	12	20	30	15	10	5

Solution:-

marks	Mid-Point (m)	Frequency (f)	$d_i = \frac{m_i - A}{h}$ $d_i = \frac{m_i - 35}{10}$	$f_i d_i$	$f_i d_i^2$	$f_i d_i^3$	$f_i d_i^4$
0-10	5	8	-3	-24	72	-216	648
10-20	15	12	-2	-24	48	-96	192
20-30	25	20	-1	-20	20	-20	20
30-40	35	30	0	0	0	0	0
40-50	45	15	1	15	15	15	15
50-60	55	10	2	20	40	80	160
60-70	65	5	3	15	45	135	405

		100		-18	240	-102	1440

Moments about origin

$$\mu_1=\frac{\sum f_i d_i}{N}\times h \qquad \Rightarrow \mu_1=\frac{-18}{100}\times 10$$

$$\mu_1=-1.8$$

$$\mu_2^{'}=\frac{\sum f_i d_i^2}{N}\times h^2 \qquad \Rightarrow \mu_2^{'}=\frac{240}{100}\times 100$$

$$\mu_2^{'}=240$$

$$\mu_3^{'}=\frac{\sum f_i d_i^3}{N}\times h^3 \qquad \Rightarrow \mu_3^{'}=\frac{-102}{100}\times 1000$$

$$\mu_3^{'}=-1020$$

$$\mu_4^{'}=\frac{\sum f_i d_i^4}{N}\times h^4 \qquad \Rightarrow \mu_4^{'}=\frac{1440}{100}\times 10000$$

$$\mu_4^{'}=144000$$

Now, we can convert moments about origin to moments about mean

$$\mu_2=\mu_2^{'}-(\mu_1^{'})^2 \qquad \Rightarrow \mu_2=240-(-1.8)^2$$

$$\mu_2=240-3.24 \qquad \Rightarrow \mu_2=236.76$$

$$\mu_3=\mu_3^{'}-3\mu_2^{'}\mu_1^{'}+2(\mu_1)^3 \qquad \Rightarrow \mu_3=-1020-3(240)(-1.8)+2(-1.8)^3$$

$$\mu_3=-1020+1296-11.664 \qquad \Rightarrow \mu_3=264.336$$

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

$$\mu_4 = 144000 - 4(-1020)(-1.8) + 6(240)(-1.8)^2 - 3(-1.8)^4$$

$$\mu_4 = 144000 - 7344 + 4665.6 - 31.4928 \qquad \Rightarrow \mu_4 = 141290.11$$

$$\beta_1 = \frac{\mu_3}{\mu_2} \qquad \Rightarrow \beta_1 = \frac{(264.336)^2}{(236.76)^3} \qquad \Rightarrow \beta_1 = 0.005$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} \qquad \Rightarrow \beta_2 = \frac{(141290.11)}{(236.76)^2} \qquad \Rightarrow \beta_2 = 2.521$$

$\therefore$ first four momemts are

$\mu'_1 = -1.8; \quad \mu'_2 = 240; \quad \mu'_3 = -1020; \quad \mu'_4 = 144000$

$\beta_1 = 0.005$ and $\beta_2 = 2.521$

[16] Compute the coefficient of skewness and kurtosis based on the following data

X:	4.5	14.5	24.5	34.5	44.5	54.5	64.5	74.5	84.5	94.5
f:	1	5	12	22	17	9	4	3	1	1

Solution:

Mid-Point (m)	Frequency (f)	$d_i = \frac{m_i - A}{h}$ $d_i = \frac{m_i - 44.5}{10}$	$f_i d_i$	$f_i d_i^2$	$f_i d_i^3$	$f_i d_i^4$
4.5	1	-4	-4	16	-64	256
14.5	5	-3	-15	45	-135	405
24.5	12	-2	-24	48	-96	192
34.5	22	-1	-22	22	-22	22
44.5	17	0	0	0	0	0
54.5	9	1	9	9	9	9
64.5	4	2	8	16	32	64
74.5	3	3	9	27	81	243
84.5	1	4	4	16	64	256
94.5	1	5	5	25	125	625
	75		-30	224	-6	2,072

$$\mu_1 = \frac{\sum f_i d_i}{N} \times h \qquad \Rightarrow \mu_1 = \frac{-30}{75} \times 10$$

$$\mu_1 = -4$$

$$\mu_2^{'} = \frac{\sum f_i d_i^2}{N} \times h^2 \qquad \Rightarrow \mu_2^{'} = \frac{224}{75} \times 100$$

$$\mu_2^{'} = 299$$

$$\mu_3^{'} = \frac{\sum f_i d_i^3}{N} \times h^3 \qquad \Rightarrow \mu_3^{'} = \frac{-6}{75} \times 1000$$

$$\mu_3^{'} = -80$$

$$\mu_4^{'} = \frac{\sum f_i d_i^4}{N} \times h^4 \qquad \Rightarrow \mu_4^{'} = \frac{1440}{100} \times 10000$$

$$\mu_4^{'} = 276300$$

$$\mu_2 = \mu_2' - (\mu_1')^2$$

$$\mu_2 = 2.99 - (-0.4)^2$$

$$\mu_2 = 2.99 - 0.16$$

$$\mu_2 = 2.83$$

$$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1')^3$$

$$\mu_3 = -0.08 + 3.588 - 0.128$$

$$\mu_3 = 3.38$$

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

$$\mu_4 = 27.63 - 4(-0.4)(-0.08) + 6(-0.4)^2(2.99) + 3(-0.4)$$

$$\mu_4 = 27.63 - 0.128 + 2.87 - 0.077$$

$$\mu_4 = 30.295$$

$$\beta_1 = \frac{\mu_3{}^2}{\mu_2{}^2} \qquad \Rightarrow \beta_1 = \frac{(3.38)^2}{(2.83)^3}$$

$$\beta_1 = \frac{11.424}{22.625} \qquad \Rightarrow \beta_1 = 0.504$$

$\beta_2 = \frac{\mu_4}{\mu_2^2}$ $\Rightarrow \beta_2 = \frac{30.295}{(2.83)^2}$

$\beta_2 = \frac{30.295}{8.01}$ $\Rightarrow \beta_2 = 3.782$

[17] The first four central moments of distribution are 0, 2.5, 0.7 and 18.75. Comment on the skewness and kurtosis as the distribution.

Solution:-we have given that

$\mu_1 = 0, \mu_2 = 2.5, \mu_3 = 0.7$ and $\mu_4 = 18.75$

skewness is measured by coefficient β_1

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

$\mu_2 = 2.5$ and $\mu_3 = 0.7$

$\therefore \beta_1 = \frac{(0.7)^2}{(2.5)^3}$ $\Rightarrow \therefore \beta_1 = 0.031$

Since $\beta_1 = 0.031$. The distribution is slightly skewed.

Testing kurtosis

$\beta_2 = \frac{\mu_4}{\mu_2^2}$ $\Rightarrow \beta_2 = \frac{18.75}{(2.5)^2}$

$\beta_2 = \frac{18.75}{6.25}$ $\Rightarrow \beta_2 = 3$

Since β_2 is exactly three, the distribution is mesokurtic

[18] Find different measures of skewness and kurtosis taking data given in example 1 of Lesson 3, using different methods.

Solution: Prepare the following table to calculate different measures of skewness and kurtosis using the values of Mean (M) = 1910, Median (M_d) = 1890.8696, Mode (M_o) = 1866.3636, Variance σ^2= 29500, Q1 = 1772.1053 and Q_3 = 2030 as calculated earlier.

Calculation of moments about an arbitrary constant 2080

Class	Mid-	freque	$X_i - A$	$f_i(X_i - A)$	$f_i(X_i - A)^2$	$f_i(X_i - A)^3$	$f_i(X_i - A)^4$

Interval	value (X_i)	ncy (f_i)					
1630-1730	1680	17	-400	-6800	2720000	-1088000000	435200000000
1730-1830	1780	19	-300	-5700	1710000	-513000000	153900000000
1830-1930	1880	23	-200	-4600	920000	-184000000	36800000000
1930-2030	1980	16	-100	-1600	160000	-16000000	1600000000
2030-2130	2080	14	0	0	0	0	0
2130-2230	2180	7	100	700	70000	7000000	700000000
2230-2330	2280	2	200	400	80000	16000000	3200000000
2330-2430	2380	2	300	600	180000	54000000	16200000000
Total		100	-40	-170	5840000	-17240	6476000000

			0	00		00000	00

First moment about the point A = 2080

$$\mu_1' = \frac{\sum_{i=1}^{n} f_i (x_i - A)^1}{N} \qquad \Rightarrow \mu_1' = \frac{-17000}{100} \qquad \Rightarrow \mu_1' = -170$$

$$\bar{x} = A + \mu_1' \qquad \Rightarrow \bar{x} = 2080 - 170 \qquad \Rightarrow \bar{x} = 1910$$

Second moment about the point A

$$\mu_2' = \frac{\sum_{i=1}^{n} f_i (x_i - A)^2}{N} \qquad \Rightarrow \mu_2' = \frac{5840000}{100}$$

$$\mu_2' = 58400$$

Third moment about the point A

$$\mu_3' = \frac{\sum_{i=1}^{n} f_i (x_i - A)^3}{N} \qquad \Rightarrow \mu_3' = \frac{-1724000000}{100}$$

$$\mu_3' = -17240000$$

Fourth moment about the point A

$$\mu_4' = \frac{\sum_{i=1}^{n} f_i (x_i - A)^4}{N} \qquad \Rightarrow \mu_4' = \frac{647600000000}{100}$$

$$\mu_4' = 6476000000$$

Compute the central moments using raw moments as follows:

$$\mu_2 = \mu_2' - (\mu_1')^2 \qquad \Rightarrow \mu_2 = 58400 - (-170)^2$$

$$\mu_2 = 29500$$

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3$$

$$\mu_3 = -17240000 - 3(58400)(-170) + 2(-170)^3$$

$$\mu_3 = 2718000$$

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

$$\mu_4 = 6476000000 - 4(-17240000)(-170) + 6(58400)(-170)^2 - 3(-170)^4$$

$$\mu_4 = 2373730000$$

Calculation of moments about mean

Class Interval	Mid-value (X_i)	Freq. (f_i)	$(x_i - \bar{x})$	$f_i(x_i - \bar{x})$	$f_i(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^3$	$f_i(x_i - \bar{x})^4$
1630-1730	1680	17	-230	-3910	899300	-206839000	47572970000
1730-1830	1780	19	-130	-2470	321100	-41743000	5426590000
1830-1930	1880	23	-30	-690	20700	-621000	18630000
1930-2030	1980	16	70	1120	78400	5488000	384160000
2030-2130	2080	14	170	2380	404600	68782000	11692940000
2130-2230	2180	7	270	1890	510300	137781000	37200870000
2230-2330	2280	2	370	740	273800	101306000	37483220000

2330-2430	2380	2	470	940	441800	207646000	97593620000
Total		100		0	2950000	271800000	237373000000

First moment about mean:

$$\mu_1 = \frac{\sum_{i=1}^{n} f_i\left(x_i - \bar{x}\right)}{N} \qquad \Rightarrow \mu_1 = \frac{0}{100}$$

$$\mu_1 = 0$$

Second moment about mean:

$$\mu_2 = \frac{\sum_{i=1}^{n} f_i\left(x_i - \bar{x}\right)^2}{N} \qquad \Rightarrow \mu_2 = \frac{2950000}{100}$$

$$\mu_2 = 29500$$

Which is equal to variance.

Third moment about mean:

$$\mu_3 = \frac{\sum_{i=1}^{n} f_i\left(x_i - \bar{x}\right)^3}{N} \qquad \Rightarrow \mu_3 = \frac{271800000}{100}$$

$$\mu_3 = 2718000$$

Fourth moment about mean:

$$\mu_4 = \frac{\sum_{i=1}^{n} f_i\left(x_i - \bar{x}\right)^4}{N} \qquad \Rightarrow \mu_4 = \frac{237373000000}{100}$$

$$\mu_4 = 2373730000$$

Karl Pearson's Coefficient of skewness (S_K):

$$S_{KP}=\frac{\text{Mean-Mode}}{\text{Standard deviation}} \qquad \Rightarrow S_{KP}=\frac{\text{M-Mo}}{\sigma}$$

$$S_{KP}=\frac{1910-1866.3636}{\sqrt{29500}} \qquad \Rightarrow S_{KP}=0.0789$$

Bowley's Coefficient of skewness (S_K):

$$S_k=\frac{Q_3+Q_1-2Q_2}{Q_3-Q_1} \qquad \Rightarrow S_k=\frac{2030+1772.1053-2\times1890.8696}{2030-1772.1053}$$

$$S_k=0.0789$$

From the above calculation the coefficient of skewness and kurtosis can be calculated as under:

$$\beta_1=\frac{\mu_3^2}{\mu_2^3} \qquad \Rightarrow \beta_1=\frac{(2718000)^2}{(29500)^3}$$

$$\beta_1=0.2878$$

$$\gamma_1=\sqrt{\beta_1} \qquad \Rightarrow \gamma_1=\sqrt{0.2878}$$

$$\gamma_1=0.5364$$

Interpretation: Hence this frequency distribution is positively skewed.

$$\beta_2=\frac{\mu_4}{\mu_2^2} \qquad \Rightarrow \beta_2=\frac{2373730000}{(29500)^2}$$

$$\beta_2=2.7276$$

$$\gamma_2=\beta_2-3 \qquad \Rightarrow \gamma_2=2.7276-3$$

$$\gamma_2=-0.2724$$

Interpretation: Hence this frequency distribution is platykurtic in nature.

Objective type questions:

(A) Fill in the blanks (Skewness)

[1] For a symmetric distribution coefficient of skewness is…

Answer:-Zero

[2] $\sqrt{\beta_1}$ and γ_1 are the **…..** coefficient of skewness.

Answer:-first

[3] Moment coefficient of skewness is called …..coefficient of skewness.

Answer:- second

[4] Measure of Central tendency which can be determined with the help of second coefficient of skewness is …….

Answer:- mode.

[5] β_1 measure the skewness but……..

Answer:- not direction

[6] Skewness means …….of the frequency distribution curve.

Answer:- asymmetry

[7] Karl Pearson's formula for measure of skewness is …..

Answer:- Coefficient of $S_K = \dfrac{\text{Mean-mode}}{\sigma}$

[8] Mean is not equal to …….in case of skew distribution.

Answer:- mode

[9] If skewness is negative then mean is ……. mode.

Answer:- less than

[10] For moderately skew distribution the relation between mean median and mode is ……..

Answer:- $(\text{mean–mode}) \cong 3(\text{mean–median})$.

[11] If a distribution has mean =7.5, mode= 10, skewness= -0.5 the variance is **……..**

Answer:-25

[12] For highly skew distribution the best measure of Central value is……..

Answer:- median

[13] Measure of skewness provides the …………of a symmetry present in distribution.

Answer:- magnitude and direction

[14] If the sum of deviations from median is not zero then a distribution will be…….

Answer:- skewed

[15] If the frequencies on either side of mode are not similarly distributed then frequency distribution curve will be……….

Answer:- skewed or asymmetrical

[16] If the mean, mode and standard deviation of a frequency distribution are 41, 45 and 8 respectively then is Pearson's coefficient of skewness is……….

Answer:- -0.5

[17] If the mean and the mode of given distribution are equal coefficient of skewness is……

Answer:- zero

[18]For a positive skewed frequency distribution the inequality that holds is ………

Answer:- $Q_3+Q_1 > 2Q_2$

[19] For a negatively skewed frequency distribution curve the third Central moment is ……….

Answer:- $\mu_3 < 0; \left[\because \sqrt{\beta_1} = \frac{\mu_3}{\sqrt{\mu_2{}^3}}\right]$

[20] For symmetrical distribution the third Central moment is...

Answer:- $\mu_3 = 0; \left[\because \sqrt{\beta_1} = \frac{\mu_3}{\sqrt{\mu_2{}^3}}\right]$

[21] In case of positive skewed distribution the relation between mean, median and mode that holds is……….

Answer:- mean > median > mode

[22] If a moderately skewed distribution has a mean 30 and mode 36 the median of the distribution is **.....**$\because \left[(\text{mean–mode}) \cong 3(\text{mean–median})\right]$

Answer:-32

[23] If a moderately skewed distribution has a mean 40 and median equal to 30 the mode of the distribution is **......**$\because \left[(\text{mean–mode}) \cong 3(\text{mean–median})\right]$

Answer:-10

[24] First and third quartiles of a frequency distribution is 30 and 75 also is coefficient of skewness is 0.6 the median of the frequency distribution is ……… $\left[\because S_{KB} = \dfrac{Q_3+Q_1-2Q_2}{Q_3-Q_1}\right]$

Answer:- 39

[25] If the mean, standard deviation and coefficient of skewness of a frequency distribution are 60 45 and -0.4 respectively the mode of the frequency distribution is….. $\left[\because \text{Coefficient of } S_K = \dfrac{\text{Mean-mode}}{\sigma}\right]$

Answer:- 78

[26] For moderately skewed distribution the empirical relation between mean median and mode is ……

Answer:- $\left[(\text{mean–mode}) \cong 3(\text{mean–median})\right]$

[27] For a negatively skewed distribution the correct relation between mean median and mode is……..

Answer:-mean $<$ median $<$ mode.

[28] In case of positive skewed distribution the extreme values lie in the …….

Answer:- Right tail.

[29] The extreme values in a negatively skewed distribution lie in the……..

Answer:- Left tail.

[30] The coefficient of skewness of series of A is 0.15 and that of series B 0.062 which of the two series is less skew ……

Answer:- Series B.

[31] For A positively skewed distribution in Karl Pearson's coefficient the inequality holds…….

Answer:- mean > mode.

[32] For A negatively skewed distribution in Karl Pearson's coefficient the inequality holds…….

Answer:- mean < mode.

[B] Fill in the blanks (Kurtosis)

[1] The relation between β_2 and γ_2 is ……

Answer:- $\gamma_2 = \beta_2 - 3$

[2] For a leptokurtic curve the relation between μ_4 and μ_2 is ….

Answer:

$$\mu_4 > 3\mu_2^2 \begin{bmatrix} \beta_2 = \dfrac{\mu_4}{\mu_2^2} & \Rightarrow \beta_2 > 3 \text{ for leptokurtic distribution} \\ \dfrac{\mu_4}{\mu_2^2} > 3 & \Rightarrow \mu_4 > 3\mu_2^2 \end{bmatrix}$$

[3] For a mesokurtic curve β_2 equal to **…….**

Answer:- three.

[4] For a platykurtic curve β_2 is less than ……

Answer:- three.

[5] For a platykurtic curve is γ_2 is ……

Answer:- less than zero.

[6] For a platykurtic curve relation between second and fourth moment is……….

Answer:- $\underline{\mu_4 < 3\mu_2^2}$

[7] Kurtosis means ……….of the frequency curve.

Answer:- bulginess

[8] If the coefficient of kurtosis is greater than 3 the distribution is …………..

Answer:- Leptokurtic.

[9] The coefficient of kurtosis is equal to 3 the frequency curve is …….

Answer:- Mesokurtic.

[10] The measure of Central tendency about kurtosis is marked is …….

Answer:- Mode.

[11] Measure of kurtosis show the degree of ……of a frequency distribution curve.

Answer:- Convexity

[12] If $\beta_2 > 3$ the distribution is said to be ……

Answer:- Leptokurtic.

[13] If the kurtosis of distribution is 3 it is called ……..distribution.

Answer:- Mesokurtic

[14] If $\beta_1 = 0$ and $\beta_2 = 3$ it is known as ……distribution curve.

Answer:- Normal

[15] standard deviation is useful in finding the descriptive measures like …………….of a distribution.

Answer:-Skewness and kurtosis

[16] If the coefficient of kurtosis of a distribution is zero then frequency is……..

Answer:-Mesokurtic $\underline{\left[\because \gamma_2 = \beta_2 - 3,\ \beta_2 = 3\right]}$

[17] If for distribution coefficient of kurtosis $\gamma_2 < 0$, the frequency curve is………

Answer:-Platykurtic $\underline{\left[\because \gamma_2 = \beta_2 - 3,\ \beta_2 < 3\right]}$

[18] Kurtosis in frequency distribution is adjusted around ………

Answer:-mode.

[19] Kurtosis and skewness of a frequency distribution are bound by the relation ……..

Answer:-their measures are always positive.

(C) Multiple choice questions: Choose the correct alternative from the following.

[1] For a negatively skewed frequency distribution curve, the third central moment is

(a) $\mu_3 > 0$ (b) $\mu_3 < 0$

(c) $\mu_3 = 0$ (d) NOTA

Answer: (b) $\mu_3 < 0$

[2] The first order moment about origin is equal to

(a) Zero (b) One

(c) Three (d) Mean

Answer: (d) Mean

[3] The first order moment about mean is equal to

(a) Zero (b) One

(c) Three (d) Mean

Answer: (a) Zero

[4] In case of symmetric distribution the odd order central moments are:

(a) Zero (b) One

(c) Positive (d) Negative

Answer: (a) Zero

[5] The central moments are invariant to the change of:

(a) Origin (b) Scale

(c) Origin and scale (d) NOTA

Answer: (a) Origin

[6] The raw moments are invariant to the change of:

(a) Origin (b) Scale

(c) Origin and scale (d) NOTA

Answer: (c) Origin and scale

[7] For a positive skewed distribution, which of the following inequality holds?

(a) Median > mode (b) Mode > mean

(c) Mean > median (d) Mean > mode

Answer: (d) Mean > mode

[8] For a negative skewed distribution, which of the following inequality holds?

(a) Median < mode (b) Mode < mean

(c) Mean < median (d) Mean < mode

Answer: (d) Mean < mode

[9] For a moderately skewed distribution has mean 30 and mode36, the median of the distribution is:

(a) 30 (b) 28

(c) 32 (d) NOTA

Answer: (c) 32

[10] For a moderately skewed distribution has mean 40 and median 30, the mode of the distribution is:

(a) 30 (b) 40

(c) 10 (d) NOTA

Answer: (c) 10

[11] If the mean, standard deviation and coefficient of skewness of a frequency distribution are 60, 45 and -0.4 respectively, the mode of the frequency distribution is:

(a) 78 (b) 45

(c) 60 (d) NOTA

Answer: (a) 78

[12] For a positive skewed distribution, which of the following relation holds?

(a) Mean < median < mode (b) Mode < median < mean

(c) Mode < mean < median (d) Mean < median < mode

Answer: (b) Mode < median < mean

[13] For a negative skewed distribution, which of the following relation holds?

(a) Mean < median < mode (b) Median < Mode < mean

(c) Mode < mean < median (d) Mean < median < mode

Answer: (d) Mean < median < mode

[14] In case of positive skewed distribution, the extreme values lie in the:

(a) Left tail (b) Right tail

(c) Middle (d) NOTA

Answer: (b) Right tail

[15] In case of negative skewed distribution, the extreme values lie in the:

(a) Left tail (b) Right tail

(c) Middle (d) NOTA

Answer: (a) Left tail

[16] The coefficient of skewness of a series A is 0.15 and that of series B is 0.062. Which of the two series is less skew?:

(a) series A (b) Series B

(c) no decision (d) NOTA

Answer: (b) Series B

[17] The coefficient of kurtosis γ_2 of a distribution is zero, the frequency curve is:

(a) Leptokurtic (b) Platykurtic

(c) Mesokurtic (d) NOTA

Answer: (c) Mesokurtic

[18] The coefficient of kurtosis β_2 of a distribution is:

(a) Less than 3 (b) Greater than 3

(c) Equal to 3 (d) NOTA

Answer: (c) Equal to 3

EXCERCISE: [D] THEORY QUESTIONS

[1] Define raw moment of frequency distribution.

[2] Define central moment of frequency distribution.

[3] Express first four central moments in terms of raw moments.

[4] Show that the central moments are invariant to the change of origin.

[5] Write a note on "Sheppard correction"

[6] What is skewness? How does it differ from dispersion? Describe the various measures of skewness.

[7] How are quartiles used for measuring dispersion and skewness?

[8] How does skewness' differ from 'dispersion' ? What is the utility of relative measures of skewness and dispersion?

[9] Explain the terms coefficient of variation, skewness and kurtosis as applied to a frequency distribution.

NUMERICAL EXAMPLES:

[1] Calculate any measures of skewness for the following:

X	0	1	2	3	4	5	6	7
F	12	27	29	19	8	4	1	0

[2] Find out standard deviation and coefficient of skewness for the following distribution:

Variable	0-5	5-10	10-15	15-20	20-25	25-30	30-35	35-40
Frequency	2	5	7	13	21	16	8	3

[3] (a) Karl Pearson's coefficient of skewness of a distribution is 0.32. Its standard deviation is 6.5 and mean is 29.6. Find the mode and median.

(b) If the mode of the above distribution is 24.8, what will be the standard deviation ?

(c) If for a distribution -0.36 is Bowley's coefficient of skewness, Q_1=8.6 and median is 12.3, find the quartile coefficient of dispersion.

[4] Particulars relating to the wage distribution of two manufacturing firms are given below:

	Firm-A Rs.	Firm-B Rs.
Mean	75	80
Median	72	70
Mode	69	62
Quartiles	62 and 78	65 and 85
Standard Deviation	13	17

Compare the features of the two distributions.

[5] (a) Compute coefficient of skewness from the following information:

Mode=18.8 inches, first quartile =14.6 and second quartile = 25.2

(b) You are given the following information:

Skewness = 0.8, Mean = 40, Mode = 36.Find the value of standard deviation.

[6](a) In a frequency distribution the coefficient of skewness based upon the quartiles is 0 6. If the sum of the upper and the lower quartiles is 100 and the median is 38, find the value of the upper quartile.

(b) A frequency distribution gives the following results:

(i) Coefficient of variation = 5 (ii) Standard deviation = 2 (ii) Karl Pearson's coefficient of skewness =0.5 Find the mean and mode of the distribution.

[7] Find the coefficient of variation of a frequency distribution, given that its mean is 120, mode is 123 and Karl Pearson's co- efficient of skewness is -0.3.

[8] For a distribution, Bowley's coefficient of skewness is -0.36, Q = 8.6 and median is 12.3. What is its quartile coefficient of dispersion?

[9] In a certain distribution the following results were obtained:-

mean = 45, median = 48, Coefficient of skewness = -0.4. Estimate the value of standard deviation with the help of the available information.

[10] The mean and standard deviation of a symmetrical distribution of weights of 12 children are 10 lbs. and 6 lbs. respectively. What is the effect on mean and standard deviation of the group if the child with the least weight and the child with the highest weight are ignored?

[11] (a) Distinguish between Pearson's and Bowley's measures of skewness.

(b) Calculate the Bowley's measures of skewness from the following:

X	10-15	15-20	20-25	25-30	30-35	35-40	40-45	45-50
f	2	5	7	13	21	16	8	3

[12] In a moderately asymmetrical distribution the mode and mean are 32.1 and 35.4 respectively. Calculate the median. Does this relationship hold good for a symmetrical distribution?