6

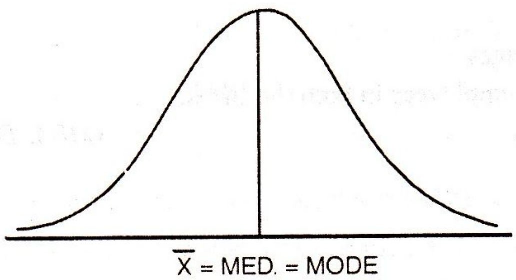
Skewness, Moments and Kurtosis

INTRODUCTION

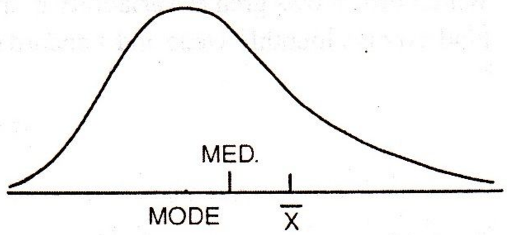
The measures of central tendency and variation discussed in previous chapters do not reveal the entire story about a frequency distribution. Two distributions may have the same mean and standard deviation but may differ in their shape of the distribution. Further description of their characteristics is necessary that is provided by measures of skewness and kurtosis.

The term 'skewness' refers to lack of symmetry or departure from symmetry, *e.g.*, when a distribution is not symmetrical (or is asymmetrical) it is called a skewed distribution. The measures of skewness indicate the difference between the manner in which the observations are distributed in a particular distribution compared with a symmetrical (or normal) distribution. 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 consequence of this assumption.

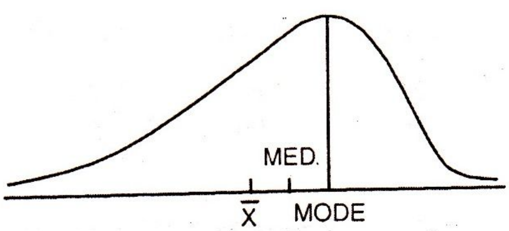
In a symmetrical distribution the values of mean, median and mode are alike. In a skewed distribution these values differ. If the value of mean is greater than the mode, skewness is said to be positive. On the other hand, if the value of mode is greater than mean, skewness is said to be negative. The following diagrams would clarify the meaning of skewness.



(*a*) Symmetrical Distribution



(*b*) Positively Skewed Distribution



(*c*) Negatively Skewed Distribution

It is clear from the (*a*), (*b*) and (*c*) diagrams that

1. In a symmetrical distribution, the values of mean, median and mode are alike.

2. In a positively skewed distribution, mean is greater than the mode and the median lies* somewhere in between mean and mode. A positively skewed distribution contains some values that are much larger than the majority of other observations.

3. In a negatively skewed distribution, mode is greater than the mean and the median lies in between mean and mode. The mean is pulled towards the low-valued item (that is, to the left). A negatively skewed distribution contains some values that are much smaller than the majority of observations.

In moderately asymmetrical distributions, the interval between the mean and the median is approximately one-third of the interval between the mean and the mode. It is this relationship that provides a means of measuring the degree of skewness.

Difference between Variation and Skewness

The following two points of difference between variation and skewness should be carefully noted :

1. Variation tells us about the amount of the variation. Skewness tells us about the direction of variation.

2. In business and economic series, measures of variation have greater practical applications than measures of skewness.

Measures of Skewness

Measures of skewness can be both absolute as well as relative. Since in a symmetrical distribution mean, median and mode are identical, the more the mean moves away from the mode, the larger the asymmetry or skewness. The distance between the mean and the mode is Karl Pearson's basis for measuring skewness. However, a measure of absolute skewness cannot be used for purposes of comparison because the same amount of skewness has different meanings in distribution with small variation and in distributions with large variation. In order to make valid comparison between the skewness in two or more distributions, we have to eliminate the distributing influence of variation. Such elimination is accomplished by dividing the absolute skewness by standard deviation. The following are two important methods of measuring relative skewness :

1. *Karl Pearson's Coefficient of Skewness.* The method is most frequently used for measuring skewness. The formula for measuring coefficient of skewness is as follows :

$$Sk_p = \frac{\text{Mean} - \text{Mode}}{\sigma}$$

Sk_p = Pearsonian (or Karl Pearson's) coefficient of skewness.

The Pearsonian coefficient of skewness is based on the same relationship as the formula for the empirical mode. The direction of skewness is determined by observing whether the mean is greater than the mode (positive skewness) or less than the mode (negative skewness). The extent of departure from symmetry is ascertained by observing the extent to which the mean is pulled away from the mode. The extent of departure is expressed in standard units in order to obtain a measure that is *independent* of the unit of measurement.

*The distance between the mode and the median is twice the distance between the median and the mean.

As the departure from symmetry becomes substantial, the relationship on which the Pearsonian coefficient formula is based breaks down and the Pearsonian coefficient no longer provides reliable results.

The value of this coefficient would be zero in a symmetrical distribution. If mean is greater than mode, coefficient of skewness would be positive, otherwise negative. In practice, the value of this coefficient usually lies between ±1 for moderately skewed distribution.

If the mode is ill-defined, the above formula has to be modified. In such a case the following approximate formula is used :

$$Sk_p = \frac{3(\text{Mean} - \text{Median})}{\sigma}$$

2. *Bowley's Coefficient of Skewness.* This method is based on quartiles. The formula for calculating coefficient of skewness is :

$$Sk_B = \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1} = \frac{Q_3 + Q_1 - 2\text{Med.}}{Q_3 - Q_1}$$

The value of this coefficient will be zero if it is a symmetrical distribution. If the value is greater than zero, it is positively skewed and if the value is less than zero, it is negatively skewed distribution.

Sk_B= Bowley's coefficient varies between ± 1 for moderately skewed distribution.

This method is particularly useful in case of open-end distributions and where extreme values are present. Also when positional measures are called for, skewness should be measured by the Bowley's method.

3. *Kelly's Coefficient of Skewness.* Another measure of skewness devised by Kelly is based on percentiles and deciles.

The formula for calculating coefficient of skewness is given below :

$$Sk_K = \frac{P_{90} - 2P_{50} + P_{10}}{P_{90} - P_{10}} \quad \text{(based on percentiles)}$$

$$Sk_K = \frac{D_9 - 2D_5 + D_1}{D_9 - D_1} \quad \text{(based on deciles)}$$

Sk_K = Kelly's coefficient of skewness.

It is clear from this formula that to calculate coefficient of skewness we have to determine the value of 10th, 50th and 90th percentiles. However, this method is not very popular in practice.

It should be noted that three different formulae of calculating skewness are based on different assumptions and hence the answer obtained from the same question by different method may differ.

It may be pointed out that measures of coefficient of skewness are used mainly for making comparison between two or more distributions. As a description of one distribution alone, the interpretation of a measure of skewness is vague as 'slight skewness', 'marked skewness', or 'moderate skewness'.

Illustration 1. The following data relate to the profits of 1,000 companies :

Profits (Rs. lakhs)	*No. of companies*	*Profits (Rs. lakhs)*	*No. of companies*
100–120	17	180–200	327
120–140	53	200–220	208
140–160	199	220–240	2
160–180	194		

Calculate the coefficient of skewness and comment on its value. *(MBA, M.D. Univ., 2001)*

Solution. CALCULATION OF COEFFICIENT OF SKEWNESS

Profits (Rs. lakhs)	*m.p. X*	*f*	*(X–170)/20 d*	*fd*	*fd²*
100–120	110	17	–3	–51	153
120–140	130	53	–2	–106	212
140–160	150	199	–1	–199	199
160–180	**170**	194	0	0	0
180–200	190	327	+1	+327	327
200–220	210	208	+2	+416	832
220–240	230	2	+3	+6	18
		$N = 1{,}000$		$\Sigma fd = 393$	$\Sigma fd^2 = 1{,}741$

$$Sk_P = \frac{\text{Mean} - \text{Mode}}{\sigma}$$

Calculation of Mean : $\bar{X} = A + \frac{\Sigma fd}{N} \times i = 170 + \frac{393}{1000} \times 20 = 170 + 7.86 = 177.86$

Calculation of Mode : By inspection mode lies in the class 180–200.

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i = 180 + \frac{133}{133+119} \times 20 = 180 + 10.56 = 190.56$$

Calculation of Standard Deviation :

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times i = \sqrt{\frac{1741}{1000} - \left(\frac{393}{1000}\right)^2} \times 20$$

$$= \sqrt{1.74 - 0.15} \times 20 = 1.26 \times 20 = 25.2$$

$$Sk_P = \frac{177.86 - 190.56}{25.2} = -0.504$$

The mode is greater than the mean by an amount equal to about 50.4 per cent of the value of standard deviation. It is a case of moderate negatively skewed distribution.

Illustration 2. The following table gives the distribution of daily wages of 500 skilled workers in a factory :

Daily wages (Rs.)	*No. of workers*
Below 200	10
200–250	25
250–300	145
300–350	220
350–400	70
400 and above	30

(*i*) Obtain the limits of daily wages of central 50 per cent of the observed workers.

(*ii*) Calculate Bowley's Coefficient of Skewness. *(MBA, Delhi Univ., 2002)*

Solution. CALCULATION OF LIMITS OF CENTRAL 50% OF WORKERS AND BOWLEY'S COEFFICIENT

Daily wages (Rs.)	*f*	*c.f.*
Below 200	10	10
200–250	25	35
250–300	145	180
300–350	220	400
350–400	70	470
400 and above	30	500

For obtaining the limits of central 50% of the workers, calculate Q_1 and Q_3

$$Q_1 = \text{Size of } \frac{N}{4}\text{th observation} = \frac{500}{4} = 125\text{th observation}$$

Q_1 lies in the class 250–300.

$$Q_1 = L + \frac{N/4 - p.c.f.}{f} \times i = 250 + \frac{125-35}{145} \times 50 = 250 + 31.03 = 281.03$$

$$Q_3 = \text{Size of } \frac{3N}{4}\text{th observation} = \frac{3 \times 500}{4} = 375\text{th observation}$$

Q_3 lies in the class 300–350.

$$Q_3 = L + \frac{3N/4 - p.c.f.}{f} \times i = 300 + \frac{375-180}{220} \times 50 = 300 + 44.32 = 344.32$$

Hence the daily wages of central 50% of workers lies between Rs. 281.03 and Rs. 344.32.

(*ii*) Bowley's Coefficient of Skewness

$$Sk_B = \frac{Q_3 + Q_1 - 2\,\text{Med.}}{Q_3 - Q_1}$$

$$\text{Med.} = \text{Size of } \frac{N}{2}\text{th observation} = \frac{500}{2} = 250\text{th observation}$$

Median lies in the class 300–350.

$$\text{Med.} = L + \frac{N/2 - p.c.f.}{f} \times i = 300 + \frac{250-180}{220} \times 50 = 300 + 15.9 = 315.9$$

$$Sk_B = \frac{344.32 + 281.03 - 2(315.9)}{344.32 - 281.03} = \frac{-6.45}{63.29} = -0.102$$

The negative coefficient –0.102 indicates that the distance between Q_3 and Q_1 is smaller than that between Q_2 and Q_1. Thus the distribution is skewed to the left or at smaller values on the X-scale.

Illustration 3. You are given the position in a factory before and after the settlement of an industrial dispute. Comment on the gains or losses from the point of view of workers and that of management :

	Before	*After*
No. of workers	3,000	2,950
Mean wage (Rs.)	2,220	2,280
Median wage (Rs.)	2,250	2,225
Standard deviation (Rs.)	300	260

Solution. The following comments can be made on the basis of information given :

(*i*) By comparing the total wage bill, we can comment on the increase or decrease in the level of wages.

	Before	*After*
Total wage bill :	3,000 × 2220 = Rs. 66,60,000	2950 × 2280 = Rs. 67,26,000

Hence the total wage bill has gone up after the settlement of dispute even though the number of workers has decreased from 3,000 to 2,950. This means that average wage is now better. This is definitely a gain to the workers. Conversely, we cannot say that increased wage bill is a loss to management because if it results in greater efficiency of workers and, therefore, higher productivity, it would be a gain to management also.

(*ii*) Median wage before settlement of the dispute was Rs. 2,250 and after settlement is Rs. 2,225. This means that formerly 50% of workers used to get wages above Rs. 2,250 and now after the settlement of dispute they get only Rs. 2,225.

(*iii*) By comparing the coefficient of variation, we can comment on the distribution pattern of wages.

	Before	*After*
Coefficient of variation :	$\frac{300}{2220} \times 100 = 13.51$	$\frac{260}{2280} \times 100 = 11.40$

Since the coefficient of variation has decreased from 13.51 to 11.40, there is sufficient evidence to conclude that wages are more uniformly distributed after the settlement of dispute, or, in other words, there is lesser inequality in the distribution of wages after the dispute is settled.

(*iv*) By comparing skewness we can comment on the nature of the distribution.

	Before	*After*
Coefficient of skewness :	$\frac{3(2220-2250)}{300} = -0.3$	$\frac{3(2280-2225)}{260} = +0.635$

The distribution was negatively skewed before the settlement and is positively skewed after the settlement.

MOMENTS

Moments are popularly used to describe the characteristic of a distribution. They represent a convenient and unifying method for summarizing many of the most commonly used descriptive statistical measures such as central tendency, variation, skewness and kurtosis. The Greek letter μ (read as mu) is generally used to denote the moments.

For Ungrouped Data

The *r*th moment of a variable *X* about the arithmetic mean $\overline{X}$ is given by :

$$\mu_r = \frac{1}{N}\Sigma\left(X-\overline{X}\right)^r \qquad ...(i)$$

The *r*th moment of a variable *X* about any arbitrary point *A* is given by :

$$\mu'_r = \frac{1}{N}\Sigma\,(X-A)^r \qquad ...(ii)$$

For Grouped Data

$$\mu_r = \frac{1}{N}\Sigma f\,(X-\overline{X})^r \qquad ...(iii)$$

and

$$\mu'_r = \frac{1}{N}\Sigma f\,(X-A)^r \qquad ...(iv)$$

For different values of *r*, we shall get different moments. Thus if we put $r = 1$, we will get first moment, if we put $r = 2$, we will get second moment, and so on.

Moments about Mean*

For ungrouped data :

$$\mu_1 = \frac{\Sigma\,(X-\overline{X})}{N}; \qquad \mu_2 = \frac{\Sigma\,(X-\overline{X})^2}{N}$$

$$\mu_3 = \frac{\Sigma\,(X-\overline{X})^3}{N}; \qquad \mu_4 = \frac{\Sigma\,(X-\overline{X})^4}{N}$$

For grouped data :

$$\mu_1 = \frac{\Sigma f\,(X-\overline{X})}{N}; \qquad \mu_2 = \frac{\Sigma f\,(X-\overline{X})^2}{N}$$

$$\mu_3 = \frac{\Sigma f\,(X-\overline{X})^3}{N}; \qquad \mu_4 = \frac{\Sigma f\,(X-\overline{X})^4}{N}$$

We can extend the moments to higher power in the similar way. But in practice the first four moments suffice.

The first moment about the origin tells us about the mean, the second moment about variance, the third moment about skewness and the fourth moment about the kurtosis.

*Moments about mean are also called central moments.

Moments about Arbitrary Point

When actual mean is in fraction, moments are first calculated about an arbitrary origin and then converted to moments about the actual mean. When deviations are taken from arbitrary point, the formulae are :

$$\mu'_1 = \frac{\Sigma (X - A)}{N} \qquad \mu'_2 = \frac{\Sigma (X - A)^2}{N}$$

$$\mu'_3 = \frac{\Sigma (X - A)^3}{N} \qquad \mu'_4 = \frac{\Sigma (X - A)^4}{N}$$

μ'_1, μ'_2, etc., denote first, second moment, etc., about an arbitrary point 'A'.

In a frequency distribution, to simplify calculations we can take a common factor but in that case the various moments have to be multiplied by i, i^2, i^3 and i^4 respectively. Thus, taking $d = \frac{X - A}{i}$ or $(X - A) = id$, we get

$$\mu'_1 = \frac{\Sigma f (X - A)}{N} \quad \text{or} \quad \frac{\Sigma fd}{N} \times i$$

$$\mu'_2 = \frac{\Sigma f (X - A)^2}{N} \quad \text{or} \quad \frac{\Sigma fd^2}{N} \times i^2$$

$$\mu'_3 = \frac{\Sigma f (X - A)^3}{N} \quad \text{or} \quad \frac{\Sigma fd^3}{N} \times i^3$$

$$\mu'_4 = \frac{\Sigma f (X - A)^4}{N} \quad \text{or} \quad \frac{\Sigma fd^4}{N} \times i^4$$

However, when we calculate the values of β_1 and β_2, the answer will remain the same whether we have multiplied the moments by common factor or not.

Finding Central Moments from Moments about Arbitrary Point

With the help of following relationships, moments about an arbitrary point can be converted to moments about mean :

$$\mu_1 = 0$$

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

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

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

Two important constants calculated from μ_2, μ_3 and μ_4 are :

(*i*) β_1 (read as beta one) and (*ii*) β_2 (read as beta two)

(*i*) β_1 is defined as : $\beta_1 = \frac{\mu_3^2}{\mu_2^3}$

β_1 is used as a measure of skewness. In a symmetrical distribution β_1 shall be zero. However, the coefficient β_1 as a measure of skewness has a serious limitation. β_1 as a measure of skewness cannot tell us about the direction of skewness, *i.e.*, whether it is positive or negative. This is for the simple reason that μ_3 being the sum of the cubes of the deviation from the mean may be positive or negative but μ_3^2 is always positive. Also μ_2 being the variance is always positive. Hence $\beta_1 = \mu_3^2/\mu_2^3$ is always positive. This drawback is removed if we calculate Karl Pearson's γ_1 (pronounced as Gamma one). γ_1 is defined as the square root of β_1, *i.e.*,

$$\gamma_1 = \sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{\mu_3}{\sigma^3}$$

The sign of skewness would depend upon the value of μ_3. If μ_3 is positive we will have positive skewness and if μ_3 is negative, we will have negative skewness.

It is advisable to use γ_1 as a measure of skewness.

(*ii*) β_2 measures kurtosis and is defined as : $\beta_2 = \frac{\mu_4}{\mu_2^2}$.

Illustration 4. From the following data calculate first four moments and also find the value of γ_1 :

Monthly Profits (in lakh Rs.)	*No. of Companies*	*Monthly Profits (in lakh Rs.)*	*No. of Companies*
Less than 7.5	4	22.5–27.5	16
7.5–12.5	10	27.5–32.5	12
12.5–17.5	20	32.5–37.5	2
17.5–22.5	36		

Solution.

CALCULATION OF MOMENTS

Monthly Profits (in lakh Rs.)	*m.p.* X	f	$(X-20)/5$ d	fd	fd^2	fd^3	fd^4
Less than 7.5	5	4	−3	−12	36	−108	324
7.5–12.5	10	10	−2	−20	40	−80	160
12.5–17.5	15	20	−1	−20	20	−20	20
17.5–22.5	20	36	0	0	0	0	0
22.5–27.5	25	16	+1	+16	16	+16	16
27.5–32.5	30	12	+2	+24	48	+96	192
32.5–37.5	35	2	+3	+6	18	+54	162
		$N = 100$		$\Sigma fd = -6$	$\Sigma fd^2 = 178$	$\Sigma fd^3 = -42$	$\Sigma fd^4 = 874$

Moments about arbitrary origin (20) in class-interval units :

$$\mu'_1 = \frac{\Sigma fd}{N} \times i = \frac{-6}{100} \times 5 = -0.3; \quad \mu'_2 = \frac{\Sigma fd^2}{N} \times i^2 = \frac{178}{100} \times 25 = 44.5;$$

$$\mu'_3 = \frac{\Sigma fd^3}{N} \times i^3 = \frac{-42}{100} \times 125 = -52.5; \quad \mu'_4 = \frac{\Sigma fd^4}{N} \times i^4 = \frac{874}{100} \times 625 = 5462.5$$

Moments about mean

$$\mu_2 = \mu'_2 - (\mu'_1)^2$$
$$= 44.5 - (-0.3)^2 = 44.5 - 0.09 = 44.41$$

$$\mu_3 = \mu'_3 - 3\mu'_1\mu'_2 + 2(\mu'_1)^3$$
$$= -52.5 - 3(-0.3 \times 44.5) + 2(-0.3)^3$$
$$= -52.5 + 40.05 - .054 = -12.504$$

$$\mu_4 = \mu'_4 - 4\mu'_1\mu'_3 + 6\mu'_2\mu'^2_1 - 3\mu'^4_1$$
$$= 5462.5 - 4(-0.3 \times -52.5) + 6(44.5)(-0.3)^2 - 3(-0.3)^4$$
$$= 5462.5 - 63 + 24.03 - .0243 = 5423.5057$$

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{-12.504}{(6.6641)^3} = -\frac{12.504}{295.954} = -0.0422 .$$

Illustration 5. The first four moments of a distribution about the value 5 of the variable are 2, 20, 40 and 50. Show that the mean is 7. Also find the other moments and β_1 and β_2.

Solution. We are given

$$\mu'_1 = 2,\ \mu'_2 = 20,\ \mu'_3 = 40 \text{ and } \mu'_4 = 50 \text{ and } A = 5.$$

We have to find the moments about mean.

$$\overline{X} = \mu'_1 + A = 2 + 5 = 7$$

$$\mu_2 = \mu'_2 - (\mu'_1)^2 = 20 - (2)^2 = 16$$

$$\mu_3 = \mu'_3 - 3\mu'_1\mu'_2 + 2\mu'^3_1 = 40 - 3\,(2)\,(20) + 2\,(2)^3 = -64$$

$$\mu_4 = \mu'_4 - 4\mu'_1\mu'_3 + 6\mu'_2\mu'^2_1 - 3\mu'^4_1 = 50 - 4\,(40)\,(2) + 6\,(20)\,(2)^2 - 3\,(2)^4 = 162$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(-64)^2}{(16)^3} = \frac{4096}{4096} = +1.00$$

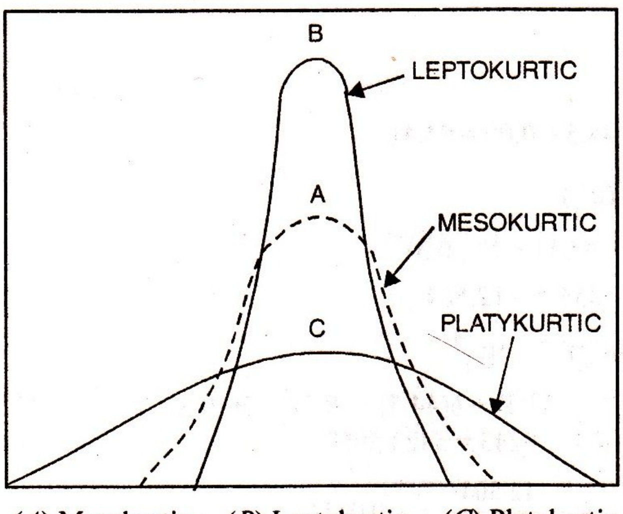
$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{162}{(16)^2} = \frac{162}{256} = +0.63$$

KURTOSIS

In describing a frequency distribution, a person can use an average to show the typical value or central tendency in the distribution, a measure of variation to show the variation of values either with certain values (such as the range and quartile deviation) or around the average of the distribution (such as the average deviation and the standard distribution) either skewed to the higher values (the right side on the *X*-scale) or to the lower values (the left side on the *X*-scale). Further, the measure of kurtosis, the fourth device in describing a frequency distribution, can be used to show the degree of concentration, either the values concentrated in the area around the mode (a peaked curve) or decentralised from the mode of both tails of the frequency curve (a flat topped curve).

Kurtosis in Greek 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 a normal curve. If a curve is more peaked than the normal curve, it is called 'leptokurtic' ; if it is more or flat-topped than the normal curve, it is called 'platykurtic' or flat-topped. The normal curve itself is known as 'mesokurtic'. The concept of kurtosis is rarely used in analysing business data :

The diagram below illustrates the scope of three different curves mentioned above :



(*A*) Mesokurtic. (*B*) Leptokurtic. (*C*) Platykurtic.

Measures of Kurtosis

Kurtosis is measured by β_2 or its derivative γ_2

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

For a symmetrical (normal) distribution the value of $\beta_2=3$ [or $\gamma_2=0$]. If the value of β_2 is greater than 3, the curve is more peaked than the normal curve; *i.e.*, leptokurtic; when the value of β_2 is less than 3, the curve is less peaked than normal curve *i.e.*, platykurtic. It may be noted that it is easier to interpret kurtosis by calculating β_2 instead of γ_2.

Illustration 6. The first central moments of a distribution are 0, 16, –36 and 120. Comment on the skewness and kurtosis of the distribution.

Solution. We are given $\mu_1 = 0$, $\mu_2 = 16$, $\mu_3 = -36$ and $\mu_4 = 120$. For commenting on the skewness we calculate γ_1.

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{-36}{(4)^3} = \frac{-36}{64} = -0.5625 \qquad \sigma = \sqrt{\mu_2} = \sqrt{16} = 4$$

The distribution is negatively skewed (It may be noted that if we calculate β_1 its value will be $\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(-36)^2}{(16)^3} = +0.3164$. But this would be wrong as μ_2 is negative). For commenting on the kurtosis we calculate β_2.

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{120}{(16)^2} = +0.469$$

Since the value of β_2 is less than 3, the distribution is platykurtic.

MISCELLANEOUS ILLUSTRATIONS

Illustration 7. An analysis of production rejects resulted in the following figures :

No. of rejects per operator	*No. of operators*	*No. of rejects per operator*	*No. of operators*
21–25	5	41–45	15
26–30	15	46–50	12
31–35	28	51–55	3
36–40	42		

Calculate mean, standard deviation and coefficient of skewness and comment on the results.

Solution. COMPUTATION OF COEFFICIENT OF SKEWNESS

No. of rejects per operator	*m.p.* *X*	*No. of operators* *f*	*(X–38)/5* *d*	*fd*	*fd²*
20.5–25.5	23	5	–3	–15	45
25.5–30.5	28	15	–2	–30	60
30.5–35.5	33	28	–1	–28	28
35.5–40.5	38	42	0	0	0
40.5–45.5	43	15	+1	+15	15
45.5–50.5	48	12	+2	+24	48
50.5–55.5	53	3	+3	+9	27
		$N = 120$		$\Sigma fd = -25$	$\Sigma fd^2 = 223$

$$\text{Mean: } \bar{X} = A + \frac{\Sigma fd}{N} \times i = 38 - \frac{25}{120} \times 5 = 38 - 1.04 = 36.96$$

$$\text{Standard deviation : } \sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times i = \sqrt{\frac{223}{120} - \left(\frac{-25}{120}\right)^2} \times 5$$

$$= \sqrt{1.8583 - .0434} \times 5 = \sqrt{1.8149} \times 5 = 1.3472 \times 5 = 6.736$$

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i = 35.5 + \frac{14}{14+27} \times 5 = 35.5 + 1.71 = 37.21$$

$$\text{Hence Coeff. of } Sk = \frac{36.96 - 37.21}{6.736} = \frac{-.25}{6.736} = -0.037$$

The value of mean = 36.96 indicates that on the average, rejects per operator were 37 in number. The value of standard deviation = 6.736 suggests that the variation in the data from the central value is approximately 7. Coefficient of skewness = –0.037 indicates that the distribution is slightly skewed to the left and therefore, there is greater concentration of the rejects per operator at the upper values than the lower values of the distribution.

Illustration 8. Distinguish between Karl Pearson's and Bowley's coefficient of skewness. Compute an appropriate measure of skewness for the following data:

Sales (Rs. Lakhs)	*No. of Companies*	*Sales (Rs. Lakhs)*	*No. of Companies*
Below 50	12	90–100	55
50–60	30	100–110	45
60–70	65	110–120	25
70–80	78	Above 120	10
80–90	80		

Solution. Since it is an open-end distribution, therefore Bowley's method of calculating skewness should be more appropriate.

CALCULATION OF COEFFICIENT OF SKEWNESS

Sales	*f*	*c.f.*
Belew 50	12	12
50–60	30	42
60–70	65	107
70–80	78	185
80–90	80	265
90–100	55	320
100–110	45	365
110–120	25	390
Above 120	10	400

$$\text{Coeff. of } Sk = \frac{Q_3 + Q_1 - 2\text{ Med.}}{Q_3 - Q_1}$$

Q_1 = Size of $\frac{N}{4}$th observation $= \frac{400}{4}$ = 100th observation.

Q_1 lies in the class 60–70.

$$Q_1 = L + \frac{N/4 - p.c.f.}{f} \times i = 60 + \frac{100 - 42}{65} \times 10 = 60 + 8.92 = 68.92$$

Q_3 = Size of $\frac{3N}{4}$th observation $= \frac{3 \times 400}{4}$ = 300th observation.

Q_3 lies in the class 90–100.

$$Q_3 = L + \frac{3N/4 - p.c.f.}{f} \times i = 90 + \frac{300 - 265}{55} \times 10 = 90 + 6.36 = 96.36$$

Med. = Size of $\frac{N}{2}$th observation $= \frac{400}{2}$ = 200th observation

Median lies in the class 80–90.

$$\text{Med.} = L + \frac{N/2 - p.c.f.}{f} \times i = 80 + \frac{200 - 185}{80} \times 10 = 80 + 1.875 = 81.875$$

$$\text{Coeff. of } Sk = \frac{96.36 + 68.92 - 2(81.875)}{96.36 - 68.92} = \frac{165.28 - 163.75}{27.44} = 0.056.$$

Illustration 9. Find an appropriate measure of skewness from the following distribution :

Age (yrs.)	*No. of employees*	*Age (yrs.)*	*No. of employees*
Below 20	13	35–40	112
20–25	29	40–45	94
25–30	46	45–50	45
30–35	60	50 and above	21

(*MBA, Bharthidasan Univ., 2007*)

Solution. Since it is an open-end distribution, therefore appropriate measure of skewness would be Bowley's coefficient of skewness.

CALCULATION OF BOWLEY'S COEFFICIENT

Age (Yrs.)	*No. of employees (f)*	*c.f.*
Below 20	13	13
20–25	29	42
25–30	46	88
30–35	60	148
35–40	112	260
40–45	94	354
45–50	45	399
50 and above	21	420
	$N = 420$	

$$Sk_B = \frac{Q_3 + Q_1 - 2\text{ Med.}}{Q_3 - Q_1}$$

$$Q_1 = \text{Size of } \frac{N}{4}\text{th observation} = \frac{420}{4} = 105\text{th observation}$$

Q_1 lies in the class 30–35.

$$Q_1 = L + \frac{N/4 - p.c.f.}{f} \times i = 30 + \frac{105 - 88}{60} \times 5 = 30 + 1.42 = 31.42$$

$$Q_3 = \text{Size of } \frac{3N}{4}\text{th observation} = \frac{3 \times 420}{4} = 315\text{th observation}$$

Q_3 lies in the class 40–45.

$$Q_3 = L + \frac{3N/4 - p.c.f.}{f} \times i = 40 + \frac{315 - 260}{94} \times 5 = 40 + 2.93 = 42.93$$

$$\text{Med.} = \text{Size of } \frac{N}{2}\text{th observation} = \frac{420}{2} = 210\text{th observation}$$

Median lies in the class 35–40.

$$\text{Med.} = L + \frac{N/2 - p.c.f.}{f} \times i = 35 + \frac{210 - 148}{112} \times 5 = 35 + 2.77 = 37.77$$

$$Sk_B = \frac{42.93 + 31.42 - (2 \times 37.77)}{42.93 - 31.42} = \frac{-1.19}{11.51} = -0.103$$

Illustration 10. (*a*) The sum of 50 observations is 500, its sum of squares is 6,000 and median 12. Find the coefficient of variation and coefficient of skewness.

Solution. $N = 50,\quad \Sigma X = 500,\quad \Sigma X^2 = 6{,}000,\quad \text{Med.} = 12$

$$\bar{X} = \frac{\Sigma X}{N} = \frac{500}{50} = 10\text{ ; and } \sigma = \sqrt{\frac{\Sigma X^2}{N} - (\bar{X})^2} = \sqrt{\frac{6{,}000}{50} - (10)^2} = 4.47$$

$$\text{C.V.} = \frac{\sigma}{\bar{X}} \times 100 = \frac{4.47}{10} \times 100 = 44.7 \text{ per cent}$$

$$\text{Mode} = 3\text{ Median} - 2\text{ Mean} = 3 \times 12 - 2 \times 10 = 16$$

$$Sk_P = \frac{\text{Mean} - \text{Mode}}{\sigma} = \frac{10 - 16}{4.47} = -1.34.$$

(*b*) For a moderately skewed distribution, the arithmetic mean is 100 and coefficient of variation is 35, and Pearson's coefficient of skewness is 0.2. Find the mode and the median.

Solution. $\bar{X} = 100$, C.V. = 35 $Sk_P = 0.2$.

$$\text{C.V.} = \frac{\sigma}{\bar{X}} \times 100$$

$$35 = \frac{\sigma}{100} \times 100 \text{ or } \sigma = 35$$

$$Sk_P = \frac{\bar{X} - \text{Mode}}{\sigma} \text{ or } 0.2 = \frac{100 - \text{Mode}}{35}$$

$$7 = 100 - \text{Mode or Mode} = 93$$

$$\text{Mode} = 3 \text{ Med.} - 2 \text{ Mean}$$

$$93 = 3 \text{ Med.} - 2 \times 100 \text{ or } 3 \text{ Med.} - 200 = 93$$

$$3 \text{ Med.} = 293 \qquad \therefore \text{ Med.} = 97.7$$

Hence Mode = 93 and Median = 97.7

Illustration 11. From the following data of age of employees, calculate coefficient of skewness and comment on the result :

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

Solution. This is a cumulative frequency distribution. First we will convert it to a simple frequency distribution and then calculate coefficient of skewness.

CALCULATION OF COEFFICIENT OF SKEWNESS

Age (Yrs.)	*m.p. X*	*f*	*(X–37.5)/5 d*	*fd*	*fd*^{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
		$N = 100$		$\Sigma fd = -5$	$\Sigma fd^2 = 275$

$$Sk_P = \frac{\text{Mean} - \text{Mode}}{\sigma}$$

Mean : $\bar{X} = A + \frac{\Sigma fd}{N} \times i = 37.5 - \frac{5}{100} \times 5 = 37.25$

Mode : Mode lies in the class 35 – 40.

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i = 35 + \frac{5}{5+10} \times 5 = 36.67$$

$$S.D. : \sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times i = \sqrt{\frac{275}{100} - \left(\frac{-5}{100}\right)^2} \times 5$$

$$= \sqrt{2.75 - .0025} \times 5 = 1.658 \times 5 = 8.29$$

$$Sk_P = \frac{37.25 - 36.67}{8.29} = \frac{0.58}{8.29} = 0.07.$$

This value of skewness indicates that the distribution has hardly any skewness.

Illustration 12. You are given the following frequency distribution of the daily earnings of employees in a company :

Earnings (in Rs.)	*Number of workers*	*Earnings (in Rs.)*	*Number of workers*
50–70	4	130–150	6
70–90	8	150–170	7
90–110	12	170–190	3
110–130	20		

Calculate the first four moments about the point 120. Convert the result into moments about the mean. Compute the value of γ_1 and γ_2 and comment on the result. (*MBA, Delhi Univ., 2002*)

Solution. Moment about some arbitrary point is given by

$$\mu_r' = \frac{1}{N} \Sigma f(X - A)^r$$

Here $A = 120$ and X are the mid-points. To get the first four moments, put $r = 1, 2, 3$ and 4 in the above formula.

COMPUTATION OF FIRST FOUR MOMENTS

Earnings (Rs.)	*m.p. X*	*f*	*(X–120)/20 d*	*fd*	*fd²*	*fd³*	*fd⁴*
50–70	60	4	–3	–12	36	–108	324
70–90	80	8	–2	–16	32	–64	128
90–110	100	12	–1	–12	12	–12	12
110–130	**120**	20	0	0	0	0	0
130–150	140	6	+1	+6	6	+6	6
150–170	160	7	+2	+14	28	+56	112
170–190	180	3	+3	+9	27	+81	243
		$N = 60$		$\Sigma fd = -11$	$\Sigma fd^2 = 141$	$\Sigma fd^3 = -41$	$\Sigma fd^4 = 825$

Moments about the arbitrary point = 120

$$\mu'_1 = \frac{\Sigma fd}{N} \times i = \frac{-11}{60} \times 20 = -3.6667$$

$$\mu'_2 = \frac{\Sigma fd^2}{N} \times i^2 = \frac{141}{60} \times (20)^2 = 940$$

$$\mu'_3 = \frac{\Sigma fd^3}{N} \times i^3 = \frac{-41}{60} \times (20)^3 = -5466.6667$$

$$\mu'_4 = \frac{\Sigma fd^4}{N} \times i^4 = \frac{825}{60} \times (20)^4 = 22{,}00{,}000$$

Moments about mean :

$\mu_1 = 0$ (since the sum of the deviations from the means is zero.)

$$\mu_2 = \mu'_2 - \mu'^2_1 = 940 - (-3.6667)^2 = 926.5553 \text{ or } \sigma = \sqrt{\mu_2} = 30.4394$$

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

$$= -5466.6667 - 3(940)(-3.6667) + 2(-3.6667)^3$$

$$= -5466.6667 + 10340.094 - 98.5953 = 4774.832$$

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

$$= 2200000 - 4(-3.6667)(-5466.6667) + 6\,(940)\,(-3.6667)^2 - 3\,(-3.6667)^4$$

$$= 2200000 - 80178.507 + 75828.045 - 542.2789 = 2195107.3$$

$$\gamma_1 = \sqrt{\beta_1} = \frac{\mu_3}{\sigma^3} = \frac{4774.83}{(30.4394)^3} = 0.1693;$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{2195107.3}{(926.56)^2} = 2.56, \quad \gamma_2 = \beta_2 - 3 = -0.44$$

The value of γ_1 indicates that the distribution is slightly skewed to the right, *i.e.*, it is not perfectly symmetrical. Since the value of γ_2 is less than zero, therefore, the distribution is platykurtic.

Illustration 13. (*a*) The first three moments of a distribution about the value 1 are 2, 25 and 80. Find its mean, standard deviation and the moment-measure of skewness.

Solution. $\mu'_1 = 2,\ \mu'_2 = 25,\ \mu'_3 = 80,\ A = 1$

Mean : $\bar{X} = \mu'_1 + A = 2 + 1 = 3$

Standard deviation : $\mu_2 = \mu'_2 - \mu'^2_1 = 25 - (2)^2 = 21$

$$\sigma = \sqrt{\mu_2} = \sqrt{21} = 4.583$$

$$\mu_3 = \mu'_3 - 3\mu'_1\mu'_2 + 2(\mu'_1)^3 = 80 - 3 \times 2 \times 25 + 2(2)^3 \quad \text{or} \quad \mu_3 = 80 - 150 + 16 = -54$$

Moment-measure of skewness : $\gamma_1 = \sqrt{\beta_1} = \dfrac{\mu_3}{\sigma^3} = \dfrac{(-54)}{(4.583)^3} = \dfrac{-54}{96.26} = -0.561$

(*b*) The first and second moment of a distribution about the value 5 of the variable are 2 and 20. Find the mean and standard deviation.

Solution.

$$\mu'_1 = 2,\ \mu'_2 = 20,\ A = 5$$

$$\overline{X} = \mu'_1 + A = 2 + 5 = 7$$

$$\mu_2 = \mu'_2 - (\mu'_1)^2 = 20 - (2)^2 = 16$$

$$\sigma = \sqrt{\mu_2} = \sqrt{16} = 4.$$

Illustration 14. Find the second, third and the fourth central moments of the frequency distribution given below. Hence find (*i*) a measure of skewness, and (*ii*) a measure of kurtosis.

Class Limits	*Frequency*	*Class Limits*	*Frequency*
110.0–114.9	5	130.0–134.9	10
115.0–119.9	15	135.0–139.9	10
120.0–124.9	20	140.0–144.9	5
125.0–129.9	35		

Solution. CALCULATION OF MOMENTS

Class Limits	*m.p.* *X*	*f*	*(X–127.45)/5* *d*	*fd*	*fd*²	*fd*³	*fd*⁴
110.0–114.9	112.45	5	–3	–15	45	–135	405
115.0–119.9	117.45	15	–2	–30	60	–120	240
120.0–124.9	122.45	20	–1	–20	20	–20	20
125.0–129.9	127.45	35	0	0	0	0	0
130.0–134.9	132.45	10	+1	+10	10	+10	10
135.0–139.9	137.45	10	+2	+20	40	+80	160
140.0–144.9	142.45	5	+3	+15	45	+135	405
		$N = 100$		$\Sigma fd = -20$	$\Sigma fd^2 = 220$	$\Sigma fd^3 = -50$	$\Sigma fd^4 = 1{,}240$

$$\mu'_1 = \frac{\Sigma fd}{N} \times i = \frac{-20}{100} \times 5 = -1; \qquad \mu'_2 = \frac{\Sigma fd^2}{N} \times i^2 = \frac{220}{100} \times 25 = 55$$

$$\mu'_3 = \frac{\Sigma fd^3}{N} \times i^3 = \frac{-50}{100} \times 125 = -62.5; \qquad \mu'_4 = \frac{\Sigma fd^4}{N} \times i^4 = \frac{1240}{100} \times 625 = 7750$$

$$\mu_2 = \mu'_2 - (\mu'_1)^2 = 55 - (-1)^2 = 55 - 1 = 54 \text{ or } \sigma = \sqrt{\mu_2} = 7.348$$

$$\mu_3 = \mu'_3 - 3\mu'_1\mu'_2 + 2\mu'^3_1 = -62.5 - 3(-1)(55) + 2(-1)^3$$

$$= -62.5 + 165 - 2 = 100.5$$

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

$$= 7750 - 4(-1)(-62.5) + 6(55)(-1)^2 - 3(-1)^4$$

$$= 7750 - 250 + 330 - 3 = 7827$$

Measure of skewness : $\gamma_1 = \dfrac{\mu_3}{\sigma^3} = \dfrac{100.5}{(7.348)^3} = \dfrac{100.5}{396.74} = +0.253$

Measure of kurtosis : $\beta_2 = \dfrac{\mu_4}{\mu_2^2} = \dfrac{7827}{(54)^2} = 2.684$

Since the value of β_2 is less than 3, the curve is platykurtic.

Illustration 15. Calculate coefficient of variation and Karl Pearson's coefficient of skewness from the data given below :

Marks		*No. of students*
Less than	30	18
" "	40	40
" "	60	70
" "	80	90
" "	100	100

(*MBA, Kumaun Univ., 2002*)

Solution.

CALCULATION OF COEFFICIENT OF VARIATION AND COEFFICIENT OF SKEWNESS

Marks	*m.p.* *X*	*f*	*(X–50)/20* *d*	*fd*	*fd*²
0 – 20	10	18	–2	–36	72
20 – 40	30	22	–1	–22	22
40 – 60	50	30	0	0	0
60 – 80	70	20	+1	+20	20
80 – 100	90	10	+2	+20	40
		N = 100		Σfd = –18	Σfd² = 154

$$\overline{X} = A + \frac{\Sigma fd}{N} \times i = 50 - \frac{18}{100} \times 20 = 50 - 3.6 = 46.4$$

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times i = \sqrt{\frac{154}{100} - \left(\frac{-18}{100}\right)^2} \times 20$$

$$= \sqrt{1.54 - 0.0324} \times 20 = 1.228 \times 20 = 24.56$$

$$\text{C.V.} = \frac{\sigma}{\overline{X}} \times 100 = \frac{24.56}{46.4} \times 100 = 52.93$$

By inspection mode lies in the class 40–60.

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i = 40 + \frac{8}{8+10} \times 20 = 40 + 8.89 = 48.89$$

$$\text{Coeff. of Sk} = \frac{\text{Mean} - \text{Mode}}{\sigma} = \frac{46.4 - 48.89}{24.56} = \frac{-2.49}{24.56} = -0.101.$$

Therefore, it is a case of low degree of negatively skewed distribution.

Illustration 16. Calculate Bowley's coefficient of skewness from the following data :

Sales (Rs. Lakhs)	*No. of Companies*
Below 50	8
" 60	20
" 70	40
" 80	65
" 90	80

(*MBA, Osmania Univ.; MBA, Delhi Univ., 2006*)

Solution.

CALCULATION OF BOWLEY'S COEFFICIENT OF SKEWNESS

Sales (Rs.Lakhs)	*No. of Companies* *f*	*c.f.*
40 – 50	8	8
50 – 60	12	20
60 – 70	20	40
70 – 80	25	65
80 – 90	15	80

$$\text{Bowley's Coeff. of Sk} = \frac{Q_3 + Q_1 - 2\text{ Med.}}{Q_3 - Q_1}$$

$$Q_1 = \text{Size of } \frac{N}{4}\text{th observation} = \frac{80}{4} = 20\text{th observation}$$

Q_1 lies in the class 50–60.

$$Q_1 = L + \frac{N/4 - p.c.f.}{f} \times i = 50 + \frac{20-8}{12} \times 10 = 50 + 10 = 60.$$

$$Q_3 = \text{Size of } \frac{3N}{4}\text{th observation} = \frac{3 \times 80}{4} = 60\text{th observation.}$$

Q_3 lies in the class 70–80.

$$Q_3 = L + \frac{3N/4 - p.c.f.}{f} \times i$$

$$= 70 + \frac{60-40}{25} \times 10 = 70 + 8 = 78$$

$$\text{Med.} = \text{Size of } \frac{N}{2}\text{th observation} = \frac{80}{2} = 40\text{th observation}$$

Median lies in the class 60–70.

$$\text{Med.} = L + \frac{N/2 - p.c.f.}{f} \times i$$

$$= 60 + \frac{40-20}{20} \times 10 = 60 + 10 = 70$$

$$\text{Coeff. of Sk} = \frac{78 + 60 - 2\,(70)}{78 - 60} = \frac{78 + 60 - 140}{18} = -0.111.$$

Therefore, it is a case of less negatively skewed distribution.

Illustration 17. The following table gives the length of life (in hours) of 400 T.V. picture tubes :

Length of life (in hours)	*No. of picture tubes*	*Length of life (in hours)*	*No. of picture tubes*
4000–4199	12	5000–5199	55
4200–4399	30	5200–5399	36
4400–4599	65	5400–5599	25
4600–4799	78	5600–5799	9
4800–4999	90		

Compute mean, standard deviation and coefficient of skewness. Comment on the values obtained. (*MBA, Delhi Univ.*)

Solution. CALCULATION OF MEAN, STANDARD DEVIATION AND COEFFICIENT OF SKEWNESS

Length of life (in hours)	*f*	*m.p.* X	*(x – 4899.5)/200* d	*fd*	*fd*²
4000–4199	12	4099.5	–4	–48	192
4200–4399	30	4299.5	–3	–90	270
4400–4599	65	4499.5	–2	–130	260
4600–4799	78	4699.5	–1	–78	78
4800–4999	90	**4899.5**	0	0	0
5000–5199	55	5099.5	+1	+55	55
5200–5399	36	5299.5	+2	+72	144
5400–5599	25	5499.5	+3	+75	225
5600–5799	9	5699.5	+4	+36	144
	$N = 400$			$\Sigma fd = -108$	$\Sigma fd^2 = 1368$

$$\bar{X} = A + \frac{\Sigma fd}{N} \times i = 4899.5 - \frac{108}{400} \times 200 = 4899.5 - 54 = 4845.5$$

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times i = \sqrt{\frac{1368}{400} - \left(\frac{-108}{400}\right)^2} \times 200$$

$$= \sqrt{3.42 - .0729} \times 200 = 1.8295 \times 200 = 365.9$$

$$\text{Coeff. of Sk} = \frac{\bar{X} - \text{Mode}}{\sigma}$$

Mode lies in the class 4800–4999. But the real limit of this class is 4799.5 – 4999.5.

$$Mode = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i = 4799.5 + \frac{12}{12+35} \times 200 = 4799.5 + 51.06 = 4850.56$$

$$\text{Coeff. of Sk} = \frac{4845.5 - 4850.56}{365.9} = \frac{-5.06}{365.9} = -0.014.$$

It is a case of very very low degree of negative skewness.

Illustration 18. You are given the following data pertaining to kilowatt hours of electricity consumed by 100 persons in Delhi :

Consumption (in K-Watt hours) :	0–10	10–20	20–30	30–40	40–50
No. of users :	6	25	36	20	13

Calculate (*i*) arithmetic mean, (*ii*) standard deviation and (*iii*) coefficient of skewness.

Solution. CALCULATION OF COEFFICIENT OF SKEWNESS

Consumption (kw. hours)	*Mid-point X*	*f*	*(X–25)/10 d*	*fd*	*fd²*
0–10	5	6	–2	–12	24
10–20	15	25	–1	–25	25
20–30	25	36	0	0	0
30–40	35	20	+1	+20	20
40–50	45	13	+2	+26	52
		N = 100		$\Sigma fd = 9$	$\Sigma fd^2 = 121$

Calculation of Mean : $\overline{X} = A + \frac{\Sigma fd}{N} \times i = 25 + \frac{9}{100} \times 10 = 25.9$

Calculation of Mode. Since the highest frequency is 36, mode lies in the class 20–30.

$$Mode = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i = 20 + \frac{11}{11+16} \times 10 = 20 + 4.07 = 24.07$$

Calculation of S.D. : $\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times i = \sqrt{\frac{121}{100} - \left(\frac{9}{100}\right)^2} \times 10$

$$= \sqrt{1.21 - .0081} \times 10 = 10.963$$

$$Sk_P = \frac{\text{Mean} - \text{Mode}}{\sigma} = \frac{25.9 - 24.07}{10.963} = 0.167.$$

Illustration 19. Calculate Karl Pearson's coefficient of skewness from the following data :

Class	*Frequency*	*Class*	*Frequency*
70–80	5	30–40	35
60–70	6	20–30	30
50–60	11	10–20	22
40–50	21	0–10	11

Solution. Arrange the class/groups and the corresponding frequencies in the ascending order.

CALCULATION OF KARL PEARSON'S COEFFICIENT OF SKEWNESS

Class	*Mid-point X*	*f*	*(X–35)/10 d*	*fd*	*fd²*
0–10	5	11	–3	–33	99
10–20	15	22	–2	–44	88
20–30	25	30	–1	–30	30
30–40	35	35	0	0	0
40–50	45	21	+1	+21	21
50–60	55	11	+2	+22	44
60–70	65	6	+3	+18	54
70–80	75	5	+4	+20	80
		N = 141		$\Sigma fd = -26$	$\Sigma fd^2 = 416$

$$\bar{X} = A + \frac{\Sigma fd}{N} \times i = 35 - \frac{26}{141} \times 10 = 35 - 1.844 = 33.156$$

Mode lies in the class 30–40.

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i = 30 + \frac{5}{5+14} \times 10 = 30 + 2.63 = 32.63$$

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times i = \sqrt{\frac{416}{141} - \left(\frac{-26}{141}\right)^2} \times 10$$

$$= \sqrt{2.95 - .034} \times 10 = 1.708 \times 10 = 17.08$$

$$Sk_p = \frac{\text{Mean} - \text{Mode}}{\sigma} = \frac{33.156 - 32.63}{17.08} = \frac{0.526}{17.08} = 0.031$$

It is a very very low degree of positive skewness.

Illustration 20. The following table gives the length of life (in hours) of 400 T.V. picture tubes:

Length of life (in hours)	*No. of picture tubes*	*Length of life (in hours)*	*No. of picture tubes*
4000–4200	22	4800–5000	80
4200–4400	38	5000–5200	70
4400–4600	65	5200–5400	50
4600–4800	75		

Compute arithmetic mean, mode, standard deviation and coefficient of skewness.

Solution. CALCULATION OF $\bar{X}$, MODE, σ AND COEFFICIENT OF SKEWNESS

Length of life (in hours)	*X*	*f*	*(X–4700)/200* *d*	*fd*	*fd²*
4000–4200	4100	22	– 3	– 66	198
4200–4400	4300	38	– 2	– 76	152
4400–4600	4500	65	– 1	– 65	65
4600–4800	4700	75	0	0	0
4800–5000	4900	80	+1	+80	80
5000–5200	5100	70	+2	+140	2s80
5200–5400	5300	50	+3	+150	450
		$N = 400$		$\Sigma fd = 163$	$\Sigma fd^2 = 1225$

Mean : $\bar{X} = A + \frac{\Sigma fd}{N} \times i = 4700 + \frac{163}{400} \times 200 = 4700 + 81.5 = 4781.5$

Mode : Mode lies in the class 4800–5000.

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i = 4800 + \frac{5}{5+10} \times 200 = 4800 + 66.67 = 4866.67$$

$$\text{S.D.} : \sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times i = \sqrt{\frac{1225}{400} - \left(\frac{163}{400}\right)^2} \times 200$$

$$= \sqrt{3.0625 - .166} \times 200 = 1.702 \times 200 = 340.4$$

$$\text{Coeff. of Sk} = \frac{\text{Mean} - \text{Mode}}{\sigma} = \frac{4781.5 - 4866.67}{340.4} = \frac{-85.17}{340.4} = -0.25.$$

Illustration 21. Calculate Karl Pearson's coefficient of skewness from the following data:

Marks	*No. of students*	*Marks*	*No. of students*
above 0	150	above 50	70
" 10	140	" 60	30
" 20	100	" 70	14
" 30	10	" 80	0
" 40	75		

(*MBA, M.D. Univ., 1998*)

Solution. This is a cumulative frequency distribution. First convert it to a simple frequency distribution and then calculate coefficient of skewness.

CALCULATION OF KARL PEARSON'S COEFFICIENT OF SKEWNESS

Marks	*m.p.* X	*f*	*(X–35)/10* d	*fd*	*fd*²	*c.f.*
0–10	5	10	– 3	– 30	90	10
10–20	15	40	– 2	– 80	160	50
20–30	25	20	– 1	– 20	20	70
30–40	35	5	0	0	0	75
40–50	45	5	+1	+5	5	80
50–60	55	40	+2	+80	160	120
60–70	65	16	+3	+48	144	136
70–80	75	14	+4	+56	224	150
		N=150		$\Sigma fd = 59$	$\Sigma fd^2 = 803$	

Since the maximum frequency 40 has been repeated twice, it is a bimodal distribution and hence we will use the formula.

$$\text{Coeff. of Sk} = \frac{3\,(\bar{X} - \text{Med.})}{\sigma}$$

$$\textit{Mean}: \ \bar{X} = A + \frac{\Sigma fd}{N} \times i = 35 + \frac{59}{150} \times 10 = 35 + 3.93 = 38.93$$

$$\textit{Median}: \text{Med.} = \text{Size of } \frac{N}{2}\text{th observation} = \frac{150}{2} = 75\text{th observation}$$

Median lies in the class 30–40

$$\text{Med.} = L + \frac{N/2 - p.c.f.}{f} \times i = 30 + \frac{75-70}{5} \times 10 = 40$$

$$\text{S.D.}: \sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times i = \sqrt{\frac{803}{150} - \left(\frac{59}{150}\right)^2} \times 10$$

$$= \sqrt{5.353 - .155} \times 10 = 2.28 \times 10 = 22.8$$

$$\text{Coeff. of Sk} = \frac{3\,(38.93 - 40)}{22.8} = \frac{3\,(-1.07)}{22.8} = \frac{-3.21}{22.8} = -0.141$$

Illustration 22. Calculate the value of γ_1 and γ_2 from the following data and interpret them:

Profits (Rs. lakhs) :	10–20	20–30	30–40	40–50	50–60
No. of Cos. :	18	20	30	22	10

Solution.

CALCULATION OF β_1 AND β_2

Profits (Rs. lakhs)	*m.p.* X	*f*	*(X–35)/10* d	*fd*	*fd*²	*fd*³	*fd*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
				$\Sigma fd = -14$	$\Sigma fd^2 = 154$	$\Sigma fd^3 = -62$	$\Sigma fd^4 = 490$

$$\mu'_1 = \frac{\Sigma fd}{N} \times i = \frac{-14}{100} \times 10 = -1.4\ ; \ \mu'_2 = \frac{\Sigma fd^2}{N} \times i^2 = \frac{154}{100} \times 100 = 154\ ;$$

$$\mu'_3 = \frac{\Sigma fd^2}{N} \times i^3 = \frac{-62}{100} \times 1000 = -620\ ; \ \mu'_4 = \frac{\Sigma fd^4}{N} \times i^4 = \frac{490}{100} \times 10000 = 49000$$

$$\mu_2 = \mu'_2 - (\mu'_1)^2 = 152.04 \text{ or } \sigma = \sqrt{\mu_2} = 12.33$$

$$\mu_3 = \mu'_3 - 3\mu'_1 \mu'_2 + 2\mu'_1{}^3 = 21.312$$

$$\mu_4 = \mu'_4 - 4\mu'_1 \mu'_3 + 6\mu'_2 (\mu'_1)^2 - (3\mu'_1)^4 = 47327.51$$

$$\gamma_1 = \frac{\mu_3}{\sigma_3} = \frac{21.312}{1874.7140} = 0.0114.$$

$$\gamma_2 = \beta_2 - 3 = \frac{\mu_4}{\mu_2{}^2} - 3 = \frac{47327.51}{(152.04)^2} - 3 = 2.047 - 3 = -0.953.$$

Therefore, $\gamma_1 = 0.0014$ suggests that it is almost near to a symmetrical distribution and γ_2 is less than zero, hence it is a platykurtic curve.

Illustration 23. Calculate Pearson's measure of skewness on the basis of mean, mode and standard deviation, from the following data :

Class-Interval :	14–15	15–16	16–17	17–18	18–19	19–20	20–21	21–22
Frequency :	35	40	48	100	125	87	43	22

(MBA, IGNOU, June 2001)

Solution : CALCULATION OF KARL PEARSON'S COEFFICIENT OF SKEWNESS

Class-Interval	*m.p.* X	*f*	*(X – 17.5)/1* d	*fd*	*fd*²
14–15	14.5	35	–3	– 105	315
15–16	15.5	40	–2	– 80	160
16–17	16.5	48	–1	– 48	48
17–18	17.5	100	0	0	0
18–19	18.5	125	+1	+ 125	125
19–20	19.5	87	+2	+ 174	348
20–21	20.5	43	+3	+ 129	387
21–22	21.5	22	+4	+ 88	352
		N = 500		Σ*fd* = 283	Σ*fd*² = 1735

$$\text{Coeff. of Sk} = \frac{\bar{X} - \text{Mode}}{\sigma}$$

Calculation of Mean : $\bar{X} = A + \frac{\Sigma fd}{N} \times i = 17.5 + \frac{283}{500} \times 1 = 17.5 + 0.57 = 18.07$

Calculation of Standard Deviation :

$$\sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times i = \sqrt{\frac{1735}{500} - \left(\frac{283}{500}\right)^2} \times i = \sqrt{3.47 - 0.32} = 1.775$$

Calculation of Mode : By inspection mode lies in the class 18–19.

$$\text{Mode} = L + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times i$$

$$L = 18,\ \Delta_1 = f_1 - f_0 = 125 - 100 = 25$$

$$\Delta_2 = f_1 - f_2 = 125 - 87 = 38,\ i = 1$$

$$\text{Mode} = 18 + \frac{25}{25 + 38} = 18 + .397 = 18.397$$

Substituting the values :

$$\text{Coeff. of Sk} = \frac{18.07 - 18.397}{1.775} = \frac{0.327}{1.775} = 0.184.$$

Illustration 24. The row data displayed below are the observations on the number of passengers who have chosen to fly on Air India in 32 cities, in a particular month.

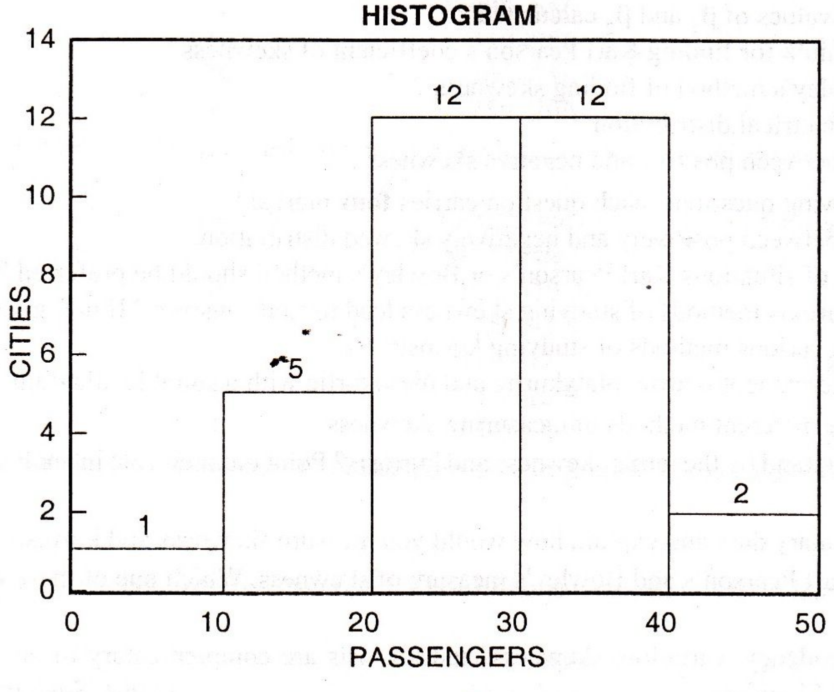
25	37	23	26	30	40	25	26
39	32	21	26	19	27	32	23
18	26	34	18	31	35	21	33
33	9	16	32	35	42	15	24

(*a*) Construct a frequency distribution using the above data.
(*b*) Develop and interpret from the above data.
(*c*) Calculate and interpret mean, median, variance and coefficient of variation for the above data.
(*d*) Are the data skewed? Give the coefficient of skewness. (*MBA, Delhi Univ., 2009*)

Solution : PREPARATION OF FREQUENCY DISTRIBUTION

Passengers	*Tally Bars*	*m.p.* m	*Cities* f	$(m-25)/10$ d	fd	fd^2	cf
0–10	|	5	1	−2	−2	4	1
10–20	𝍸	15	5	−1	−5	5	6
20–30	𝍸 𝍸 ||	25	12	0	0	0	18
30–40	𝍸 𝍸 ||	35	12	+1	+12	12	30
40–50	||	45	2	+2	+4	8	32
			$N = 32$		$\Sigma fd = 9$	$\Sigma fd^2 = 29$	

HISTOGRAM



$$\text{Mean}: \overline{X} = A + \frac{\Sigma fd}{N} \times i = 25 + \frac{9}{32} \times 10 = 25 + 2.813 = 27.813$$

$$\text{Median}: \text{Med.} = \text{Size of } \frac{N}{2}\text{th item} = \frac{32}{2} = 16\text{th item}$$

Median lies in the class 20–30

$$\text{Med.} = L + \frac{N/2 - p.c.f.}{f} \times i$$

$$= 20 + \frac{16-6}{12} \times 10 = 20 + 8.33 = 28.33$$

$$\text{Standard Deviation}: \sigma = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} \times i$$

$$= \sqrt{\frac{29}{32} - \left(\frac{9}{32}\right)^2} \times 10 = \sqrt{0.906 - 0.079} \times 10$$

$$= 0.909 \times 10 = 0.09$$

$$\text{Variance}: \sigma^2 = (9.09)^2 = 82.623$$

$$\text{Coeff. of Variation} = \frac{\sigma}{\overline{X}} \times 100 = \frac{9.09}{27.813} \times 100 = 32.68$$

Since it is a bi modal series, skewness will be calculated by formula :

$$\text{Coeff. of Sk} = \frac{3(\overline{X} - \text{Med.})}{\sigma}$$

$$\overline{X} = 27.813, \quad \text{Median} = 28.33, \quad \sigma = 9.09$$

$$\text{Coeff. of Sk} = \frac{3(27.813 - 28.33)}{9.09} = \frac{-1.551}{9.09} = -0.171$$

The distribution is skewed to the left. However, there is very low degree of skewness.

PROBLEMS

1-A: Answer the following questions, each question carries **one** mark:

(*i*) What is skewness ?
(*ii*) Point out the role of studying skewness.
(*iii*) Name the various methods of finding skewness.
(*iv*) What are kurtosis ?
(*v*) What are moments ?
(*vi*) How are the values of β_1 and β_2 calculated ?
(*vii*) Give the formula for finding Karl Pearson's coefficient of skewness.
(*viii*) What is Bowley's method of finding skewness ?
(*ix*) What is symmetrical distribution ?
(*x*) Distinguish between positive and negative skewness .

1-B: Answer the following questions, each question carries **four** marks:

(*i*) Distinguish between positively and negatively skewed distribution.
(*ii*) In what type of situations Karl Pearson's or Bowley's method should be preferred ?
(*iii*) Would the various methods of studying skewness lead to same answer ? If not, give reasons.
(*iv*) What are the various methods of studying kurtosis ?
(*v*) Explain the terms leptokurtic, platykurtic and mesokurtic with a suitable diagram.

2. (*a*) Explain briefly the different methods of measuring skewness.
(*b*) What do you understand by the terms skewness and kurtosis? Point out their role in analysing a frequency distributi
(MBA, Delhi Univ., 20

3. Take any suitable imaginary data and explain how would you measure skewness and kurtosis.

4. Distinguish between Karl Pearson's and Bowley's measure of skewness. Which one of these would you prefer and wh
(MBA, Delhi Univ., 2

5. Measures of central, tendency, variation, skewness, and kurtosis are complementary to one another in understandi frequency distribution? Elucidate. *(MBA, Sukhadia Univ.; Delhi Univ., 2*

6. Define 'Moments'. How can you find out skewness and kurtosis of a distribution from moments about the mean?

7. Explain clearly how the moments help in describing the characteristics of a frequency distribution.

8. Explain clearly how the measures of skewness and kurtosis can be used in describing a frequency distribution.

9. What is meant by 'moments' of a distribution ? Show how moments are used to describe the characteristics of a distrib *i.e.*, central tendency, dispersion, skewness and kurtosis.

10. What are the raw and the central moments of a distribution? Show that the central moments are invariant under cha origin but not under change of scale.

11. Define raw and central moments of a frequency distribution. Express the second, third and fourth central moments i of raw moments.

12. (*a*) Explain the terms 'Skewness' and 'Kurtosis' used in connection with the frequency distribution of a continuous v Give the different measures of skewness (any two of the measures to be given) and kurtosis.
(*b*) Define and discuss the 'quartiles' of a distribution. How are they used for measuring variation and skewness

13. Define moments. Establish the relationship between the moments about mean in terms of moments about any arbitrary point and *vice-versa*.

14. (*a*) Define moments. How are they helpful in study of the different aspects of the formation of a frequency distribution?
(*b*) "A frequency distribution can be described almost completely by the first four moments and the two measures based on the moments." Examine.

15. (*a*) Explain the third and fourth central moment in terms of the first four moments about the origin.
(*b*) Distinguish between variation and skewness and point out the various methods of measuring skewness.
(*c*) Explain the term 'skewness'. What purpose does a measure of skewness serve? Comment on some of the well-known measures of skewness.

16. (*a*) Distinguish between skewness and kurtosis.
(*b*) Briefly mention the tests which can be applied to determine the presence of skewness.

17. (*a*) How do measures of central tendency, dispersion, skewness and kurtosis help in analysing a frequency distribution? Explain with the help of an example. (*MBA, Sukhadia Univ., 2008*)
(*b*) Find out coefficient of skewness from the following table giving wages of 240 persons :

Wages (Rs.)	*No. of persons*	*Wages (Rs.)*	*No. of persons*
2000–2200	12	2800–3000	50
2200–2400	18	3000–3200	45
2400–2600	35	3200–3400	30
2600–2800	42	3400–3600	8

[Sk = – 0.267]

18. Calculate Karl Pearson's coefficient of skewness from the following data :

Profits (Rs. Lakhs)	*No. of Cos.*	*Profits (Rs. Lakhs)*	*No. of Cos.*
400–450	8	600–650	62
450–500	10	650–700	32
500–550	30	700–750	15
550–600	45	750–800	8

19. The following data represent the percentage of ash content in a particular variety of coal as determined by test on 280 wagon loads :

Percentage of ash content	*Frequency*	*Percentage of ash content*	*Frequency*
Less than 6.0	0	10.0 –10.9	84
6.0– 6.9	1	11.0 –11.9	45
7.0– 7.9	7	12.0–12.9	28
8.0– 8.9	28	13.0–13.9	7
9.0– 9.9	78	14.0–14.9	2

Calculate the quartile coefficient of skewness. Also compare the proportion of the total frequency outside the limits $\overline{X} \pm 2\sigma$ for the distribution.

[Sk=0.05; 2.3]

20. From the following data of daily travelling allowance (in Rs.) of salesmen, calculate coefficient of skewness and comment on its value :

Travelling allowance (per day)	*No. of salesmen*	*Travelling allowance (per day)*	*No. of salesmen*
110–115	4	135–140	90
115–120	10	140–145	52
120–125	26	145–150	33
125–130	49	150–155	17
130–135	72	155–160	7

21. From the following data pertaining to profits (Rs. lakhs) for 50 companies, calculate moments β_1 and β_2 :

Profits (Rs. Lakhs)	*No. of Compaines*
70–90	8
90–110	11
110–130	18
130–150	9
150–170	4

[$\mu_2 = 528$, $\mu_3 = 960$, $\mu_4 = 642816$, $\beta_1 = 0.006$, $\beta_2 = 2.31$]

22. A record was kept over a period of 6 months by a sales manager to determine the average number of calls made per day by his six salesmen. The results are shown below :

Salesmen	:	A	B	C	D	E	F
Average number of calls per day	:	8	10	12	15	7	5

(i) Compute a measure of skewness. Is the distribution symmetrical ?
(ii) Compute a measure of kurtosis. What does this measure mean ?
[β_1=0.11; β_2=1.97]

23. Locate the mode and calculate mean and standard deviation of the following distribution and using your results comment on the skewness of the distribution :

Scores	Frequency	Scores	Frequency
10–15	2	35–40	6
15–20	8	40–45	4
20–25	6	45–50	3
25–30	12	50–55	1
30–35	7	55–60	1

[$\bar{X}$ = 30.1; Mo.= 27.73, σ = 10.45, Sk = 0.227]. *(MBA, Delhi Univ., 2002, 2005)*

24. You are given the following information before and after the settlement of an industrial dispute :

	Before settlement of dispute	After settlement of dispute
No. of workers	1100	950
Average wage (Rs.)	2350	2400
Standard deviation (Rs.)	425	400
Median wage (Rs.)	2375	2325

Comment on the gains and losses from the point of view of workers and that of management.

25. The arithmetic mean of a distribution is 5. The second and the third moments about the mean are 20 and 140 respectivel Find the third moment of the distribution about 10.
[–285]

26. For the frequency distribution given below, calculate the coefficient of skewness based on the quartiles :

Class limits	Frequency	Class limits	Frequency
10–19	5	50–59	25
20–29	9	60–69	5
30–39	14	70–79	8
40–49	20	80–89	4

27. (a) For a distribution, Bowley's coefficient of skewness is –0.48, Q_3 = 10.2 and Median = 14.4. What is the quarti coefficient of distribution?
(b) Karl Pearson's coefficient of skewness of a distribution is +0.4. Its standard deviation is 10 and mean 40.5. Find tl mode and median of the distribution.
(c) Find coefficient of skewness from the information given below :

$$Q_1 = 60,\ Q_3 = 75,\ \text{Med.} = 68.$$

(d) The following information was obtained from the records of a factory relating to wages; $\bar{X}$ = 275, Med. = 260, σ 45.8
Give as much information as you can about the distribution of wages.
[(a) 0.22 (b) 39.17 (c) –0.07 (d) Sk = 0.98]

28. The first three moments of a distribution about the value 7 calculated from a set of 9 observations are 0.2, 19.4 and –41 Find the measures of central tendency and dispersion and also the third moment about origin.
[7.2, 4.4, –52.624]

29. The first four moments of a distribution about A = 4 are 1, 4, 10 and 45. Obtain the various characteristics of the distributi on the basis of the information given. Comment upon the nature of the distribution.
[β_1 = 0, β_2 = 2.897]

30. (a) State the use of quartiles for measuring dispersion and skewness.
(b) Calculate Bowley's coefficient of skewness from the following data :

Mid-value :	75	100	125	150	175	200	225	250
Frequency :	35	40	48	100	125	80	50	22

[–0.032]

31. A prospective buyer tested the bursting pressure of the sample of polythene bags received from a manufacturer. The test gives the following results :

Bursting pressure (in lbs.)	5–10	10–15	15–20	20–25	25–30	30–35
No. of bags	2	20	30	50	6	2

The buyer calculated the mean and mode of the sample as 20.2 lbs. and 21.5 lbs. respectively. Calculate (*i*) coefficient of variation, (*ii*) Karl Pearson's coefficient of skewness for bursting pressure.

32. From the following data, calculate coefficient of variation and coefficient of skewness :

Age (in years) :	25–30	30–35	35–40	40–45	45–50	50–55	55–60
No. of employees :	9	18	30	40	10	7	6

33. The frequency distribution of weekly wages (in Rs.) in a certain factory is as follows :

Weekly wages	*No. of workers*	*Weekly wages*	*No. of workers*
423–427	2	448–452	16
428–432	6	453–457	12
433–437	9	458–462	6
438–442	14	463–467	2
443–447	32	468–472	1

Find Karl Pearson's coefficient of skewness and interpret its value.
[$Sk_P = 0.0572$]

34. A survey was conducted by a manufacturing company to enquire the maximum price at which persons would be willing to buy their product. The following table gives the stated prices (in rupees) by persons :

Price (in Rs.) :	80–90	90–100	100–110	110–120	120–130
No. of persons :	11	29	18	27	15

Calculate Bowley's coefficient of skewness and interpret its value. *(MBA, Delhi Univ., 2002)*

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

36. Calculate coefficient of variation and Karl Pearson's coefficient of skewness from the data given below :

Sales (Rs. crores) Less than :	40	50	60	70	80
No. of Companies :	8	20	50	72	80

[Coeff. of Variation = 19.55, Coeff. of Sk = –0.06]

37. Assume that a firm has selected a random sample of 100 from its production line and has obtained the data shown in the table below :

Class-interval	*Frequency*	*Class-interval*	*Frequency*
130–134	3	150–154	19
135–139	12	155–159	12
140–144	21	160–164	5
145–149	28	Total	100

Compute Karl Pearson's Coefficient of Skewness.
[Coeff. of Sk = –0.572]
(MBA, Mangalore Univ., 2005)

38. (*a*) A moderately skewed distribution has mean and median as 25 and 26 respectively. Then its mode approximately equals......
(*b*) Whether the following statement is *true* or *false* : If a distribution has negative skewness then its mean is greater than mode.

39. Calculate the first four moments about mean and find the values of β_1 and β_2 and comment on the result :

Profits (Rs. lakhs) :	0–10	10–20	20–30	30–40	40–50	50–60	60–70
No. of Companies :	8	12	20	30	15	10	5

(MBA, Kumaun Univ., 2004)

40. From the following data pertaining to the income of 5,800 persons, find Bowley's coefficient of skewness and interpret its value :

Income (Rs.)	*No. of persons*	*Income (Rs.)*	*No. of persons*
Below 10,000	170	40,000–50,000	1,350
10,000–20,000	630	50,000–60,000	1,000
20,000–30,000	1,000	60,000 and above	400
30,000–40,000	1,250		

[Coeff. of Sk = –0.067]

(MBA, Kurukshetra Univ., 2001)

41. Compute the first 3 moments about the arithmetic mean from the following data :

Variable value :	5	10	15	20	25	30	35
Frequency :	8	15	20	32	23	17	5

(*MBA, Lucknow Univ., 2001*)

42. The following distribution gives the pattern of overtime work done in a month by 100 employees of a company :

Overtime hours :	10–15	15–20	20–25	25–30	30–35	35–40
No. of employees :	11	20	35	20	8	6

Compute mean, mode, standard deviation and coefficient of skewness.
[23.1, 22.5, 6.4915, 0.0924]

43. The following table gives the distribution of monthly wages of 500 workers in a factory:

Monthly wages (Rs. hundred)	*No. of workers*	*Monthly wages (Rs. hundred)*	*No. of workers*
15–20	10	30–35	220
20–25	25	35–40	70
25–30	145	40–45	30

Compute Karl Pearson's and Bowley's coefficient of skewness. Interpret the values.
[$SK_P = -0.022$, $SK_B = -0.102$] (*MBA, Delhi Univ., 2006*)

44. Calculate Karl Pearson's coefficient of skewness from the data given below :

Marks	*No. of candidates*	*Marks*	*No. of candidates*
70–80	11	30–40	21
60–70	22	20–30	11
50–60	30	10–20	6
40–50	35	0–10	5

[–0.026] (*MBA, Kumaun Univ., 2001*)

45. Calculate β_1 and β_2 from the following distribution and interpret the results :

Age	*Frequency*	*Age*	*Frequency*
25–30	2	45–50	25
30–35	8	50–55	16
35–40	18	55–60	7
40–45	27	60–65	2

[$\beta_1 = 0.034$, $\beta_2 = 2.59$]

*****