'Measures of skewness indicates the tendency of the group to depart from symmetry'

– Connor

5

Skewness of Frequency Distribution

Contents :

5.1 Meaning of Skewness

5.2 Types of Skewness

5.3 Concept of Absolute and Relative Measures of Skewness

5.4 Methods of obtaining Measures of Skewness and Coefficients of Skewness

5.4.1 Karl Pearson's Method

5.4.2 Bowley's Method

5.5 Comparison of two methods of Coefficient of Skewness

5.1 Meaning of Skewness

Measures of central tendency and dispersion give important information about the population under study. These measures give the information about the values assumed by the units of the population. Moreover, a comparison of two or more populations and an analysis to obtain other information is possible using these measures. We have studied the nature of population observations by these measures.

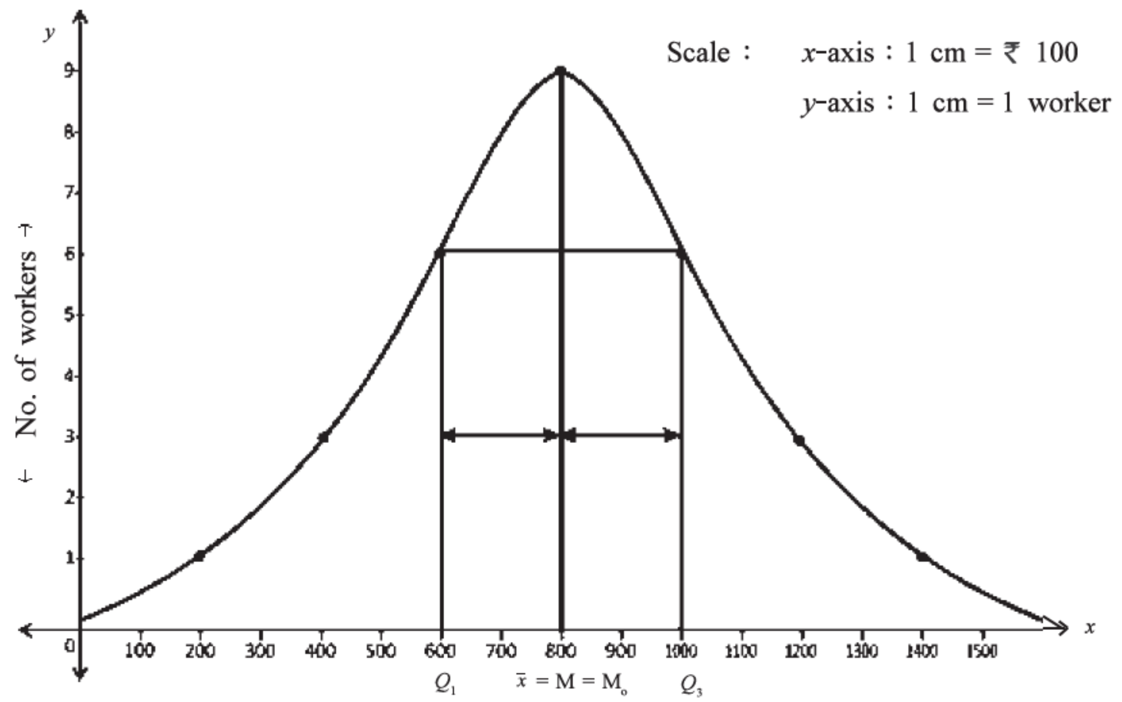
The measures of central tendency and dispersion give us partial information of the central tendency and dispersion of observations around the central tendency measure of raw data. But we cannot get its complete information. To get more information about this, the population is represented by frequency curve. Thus, the direction, shape and form of the frequency curve can be studied from the frequency curve of the frequency distribution. We will study the third important measure called skewness to obtain more information about the population. Let us understand the concept of symmetric frequency distribution before explaining the meaning of skewness.

Symmetric Frequency Distribution

A frequency distribution in which the observations of the population are evenly distributed on both the sides of the mode is called symmetric frequency distribution and its frequency curve is called symmetric frequency curve. Generally, the frequency curve of symmetric distribution is found to be bell-shaped. We shall study the concept of symmetry by the frequency curve and measures of central tendency using the illustration of frequency distribution of daily wages (in ₹) of workers.

Daily wage of worker (₹)	200	400	600	800	1000	1200	1400
No. of workers	1	3	6	9	6	3	1

To ascertain whether the above frequency distribution is symmetric or not, we will draw the frequency curve and find the measures of central tendency.



← daily wage of worker (₹) →

Thus, the frequency curve of the given distribution is bell-shaped. The observation ₹ 800 has maximum frequency 9. The observations at equal distance on both the sides of ₹ 800 have the same frequencies. Hence, the observations on both the sides of mode of the distribution are equally distributed.

Now we will compute mean, median, mode and quartiles of the frequency distribution and use them to understand the symmetric distribution.

Daily wage of workers (₹) x	200	400	600	800	1000	1200	1400	Total
No. of workers f	1	3	6	9	6	3	1	n = 29
fx	200	1200	3600	7200	6000	3600	1400	23,200
cf	1	4	10	19	25	28	29	

Mean $\bar{x} = \frac{\Sigma fx}{n}$

$= \frac{23200}{29}$

$= 800$

$\therefore \quad \bar{x} = ₹\ 800$

Median M = Value of the $\left(\frac{n+1}{2}\right)$th observation

= Value of the $\left(\frac{29+1}{2}\right)$th observation

= Value of the $\left(\frac{30}{2}\right)$th observation

= Value of the 15th observation

Referring to *cf* column, we find that the value of the 15th observation is 800.

$\therefore \quad M = 800$

$\therefore \quad M = ₹\ 800$

Mode M_o = observation with highest frequency 9 = 800

$\therefore \quad M_o = ₹\ 800$

First quartile Q_1 = Value of the $\left(\frac{n+1}{4}\right)$th observation

= Value of the $\left(\frac{29+1}{4}\right)$th observation

= Value of the 7.5th observation

Referring to *cf* column, we find that the value of the 7.5th observation is 600

$\therefore \quad Q_1 = 600$

$\therefore \quad Q_1 = ₹\ 600$

Third quartile Q_3 = Value of the $3\left(\frac{n+1}{4}\right)$th observation

= Value of the $3\left(\frac{29+1}{4}\right)$th observation

= Value of the 3(7.5)th observation

= Value of the 22.5th observation

Referring to *cf* column, we find that the value of the 22.5th observation is 1000.

$\therefore \quad Q_3 = 1000$

$\therefore \quad Q_3 = ₹\ 1000$

It is clear from the values of mean, median, mode and quartiles of the distribution as well as the frequency curve of the distribution that,

(1) We have $\bar{x} = M = M_o = ₹\ 800$

(2) $(Q_3 - M) = (M - Q_1)$

$\therefore \quad (1000 - 800) = (800 - 600)$

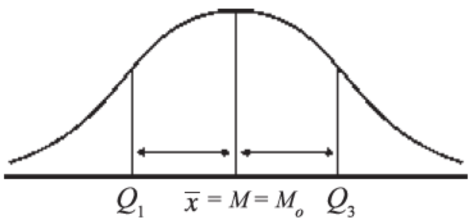
Thus, the quartiles are equidistant from median.

(3) The frequency curve of the frequency distribution is found to be bell-shaped.

Thus, **the following characteristics are generally observed in a symmetric frequency distribution.**

(i) Mean, median and mode have same value.
That is, $\bar{x} = M = M_o$

(ii) First quartile Q_1 and third quartile Q_3 are at equal distance from median M.
That is, $(Q_3 - M) = (M - Q_1)$



(iii) The frequency curve of the frequency distribution is found to be bell-shaped.

(iv) The frequency of observation at equal distance on both the sides of mode is equally distributed.

If the above characteristics are absent in any given frequency distribution then it is said that the frequency distribution is not symmetric. **The frequency distribution with lack of symmetry is called skewed frequency distribution. Thus, lack of symmetry is called skewness.** The following situations indicate skewness in a frequency distribution :

(1) The values of mean, median and mode are not same.

(2) Quartiles Q_1 and Q_3 are not at equal distances from median M (= Q_2).

Thus $(Q_3 - M) \neq (M - Q_1)$

(3) The right or left tail of the frequency curve is more elongated.

(4) The frequency of observation at equal distance on both sides of mode is not equally distributed.

The above conditions are called tests of skewness as they check whether the frequency distribution is skewed or not. Now we shall study the types of skewness.

5.2 Types of Skewness

There are two types of skewness for frequency distribution: (l) Positive Skewness and (2) Negative Skewness. We understand these two types with diagram and illustration.

(1) Positive Skewness :

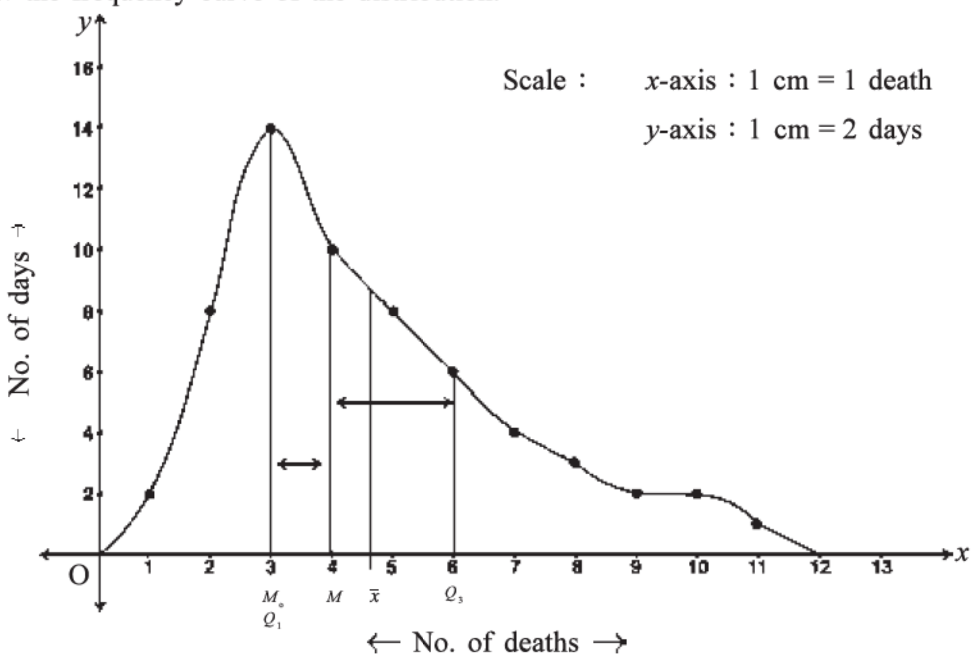
If the right tail of the frequency curve of a distribution is more elongated then it is called positively skewed distribution. The population is said to possess positive skewness due to this characteristic.

Now we take the following illustration to study positive skewness using frequency curve and measures of central tendency measures.

The frequency distribution of number of deaths in a hospital of a city during 60 days is as follows :

No. of deaths	1	2	3	4	5	6	7	8	9	10	11
No. of days	2	8	14	10	8	6	4	3	2	2	1

We will draw the frequency curve of the distribution.



The right tail of the frequency curve is elongated. Less observations are distributed to the left of the observation $x = 3$ corresponding to the maximum frequency 14. More observations are distributed to its right side and their frequencies are decreasing gradually.

Now we find the mean, median, median and quartiles of the distribution.

No. of deaths x	No. of days f	fx	cf
1	2	2	2
2	8	16	10
3	14	42	24
4	10	40	34
5	8	40	42
6	6	36	48
7	4	28	52
8	3	24	55
9	2	18	57
10	2	20	59
11	1	11	60
Total	$n = 60$	**277**	

Mean $\bar{x} = \frac{\Sigma fx}{n}$

$= \frac{277}{60}$

$= 4.6166$

$\therefore \quad \bar{x} \doteqdot 4.62$ deaths

Mode M_o = Observation with highest frequency 14 = 3

$\therefore \quad M_o$ = 3 deaths

Median M = Value of the $\left(\frac{n+1}{2}\right)$th observation

= Value of the $\left(\frac{60+1}{2}\right)$th observation

= Value of the $\left(\frac{61}{2}\right)$th observation

= Value of the 30.5th observation

Referring to cf column, we find that the value of the 30.5th observation is 4.

$\therefore \quad M$ = 4 deaths

First quartile Q_1 = Value of the $\left(\frac{n+1}{4}\right)$th observation

= Value of the $\left(\frac{60+1}{4}\right)$th observation

= Value of the $\left(\frac{61}{4}\right)$th observation

= Value of the 15.25th observation

Referring to cf column, we find that the value of the 15.25th observation is 3.

$\therefore \quad Q_1$ = 3 deaths

Third quartile Q_3 = Value of the $3\left(\frac{n+1}{4}\right)$th observation

= Value of the $3\left(\frac{60+1}{4}\right)$th observation

= Value of the 3(15.25)th observation

= Value of the 45.75th observation

Referring to *cf* column, we find that the value of the 45.75th observation is 6.

$\therefore \quad Q_3 = 6$ deaths

Thus, the following results are obtained for this frequency distribution :

(1) $\bar{x} = 4.62$, $M = 4$ and $M_o = 3$ Hence, $\bar{x} > M > M_o$

(2) $Q_3 - M = 6 - 4 = 2$ and $M - Q_1 = 4 - 3 = 1$. Hence $Q_3 - M > M - Q_1$

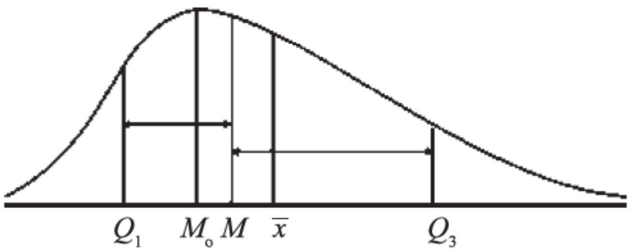
(3) The right tail of the frequency curve is more elongated.

Thus the following characteristics are generally observed in a positively skewed frequency distribution :

(1) The values of median and mode are in decreasing order in this distribution. That is, $\bar{x} > M > M_o$

(2) The distance between third quartile Q_3 and median M is more than the distance between median and first quartile Q_1. That is $(Q_3 - M) > (M - Q_1)$

(3) The right tail of the frequency curve of this distribution is more elongated.



Note : We find more variation in the observations to the right of the mode in a positively skewed distribution. e.g. positive skewness is found in the frequency distribution of number of deaths.

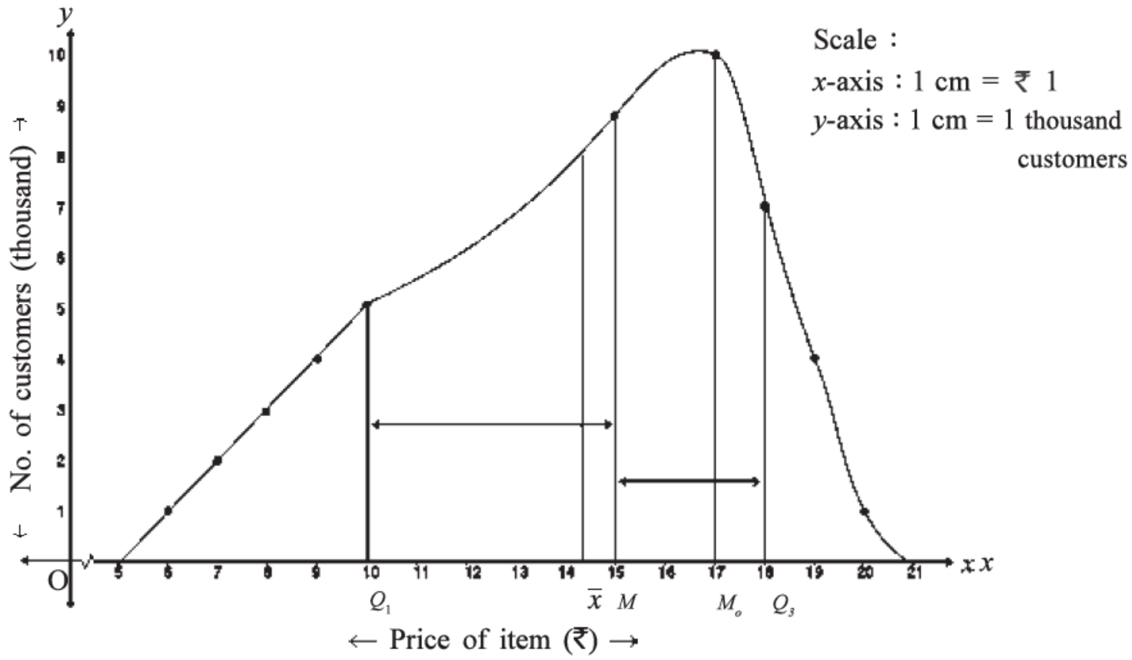
(2) Negative Skewness

If the left tail of the frequency curve of a distribution is more elongated then it is called negatively skewed distribution. Such a population is said to possess negative skewness.

We will study the negative skewness obtaining the frequency curve and central tendency measures of the following frequency distribution.

Price of an item (₹)	6	7	8	9	10	15	17	18	19	20	Total
No. of customers (thousand)	1	2	3	4	5	8	10	7	4	1	45

We will draw the frequency curve of this distribution.



The left tail of the frequency curve is elongated. Less observations are distributed to the right of observation $x = 17$ corresponding to maximum frequency 10 whereas more observations are distributed to the its left and their frequencies are decreasing gradually.

Now we find mean, median, mode and quartiles for the distribution.

Price of item (₹) x	**No. of customers f**	fx	cf
6	1	6	1
7	2	14	3
8	3	24	6
9	4	36	10
10	5	50	15
15	8	120	23
17	10	170	33
18	7	126	40
19	4	76	44
20	1	20	45
Total	$n = 45$	**642**	

Mean $\bar{x} = \frac{\Sigma fx}{n}$

$= \frac{642}{45}$

$= 14.27$

$\bar{x} = ₹\ 14.27$

Median M = value of the $\left(\frac{n+1}{2}\right)$th observation

= value of the $\left(\frac{45+1}{2}\right)$th observation

= value of the $\left(\frac{46}{2}\right)$th observation

= value of the 23rd observation

Reffering to *cf* column, we find that the value of the 23rd observation is 15.

$\therefore$ M = ₹ 15

First quartile Q_1 = value of the $\left(\frac{n+1}{4}\right)$th observation

= value of the $\left(\frac{45+1}{4}\right)$th observation

= value of the $\left(\frac{46}{4}\right)$th observation

= value of the 11.5th observation

Reffering to *cf* column, we find that the value of the 11.5th observation is 10.

$\therefore$ Q_1 = ₹ 10

Third quartile Q_3 = value of the $3\left(\frac{n+1}{4}\right)$th observation

= value of the $3\left(\frac{45+1}{4}\right)$th observation

= value of the $3\left(\frac{46}{4}\right)$th observation

= value of the 3(11.5)th observation

= value of the 34.5th observation

Reffering to *cf* column, we find that the value of the 34.5th observation is 18.

$\therefore$ Q_3 = ₹ 18

Mode M_o = observation with highest frequency 10 = 17

$\therefore$ M_o = ₹ 17

From the frequency curve showing mean $\bar{x}$, median M, mode M_o and the values of quartiles it is clear that

(1) $\bar{x} = 14.27$, $M = 15$, $M_o = 17$ Hence $\bar{x} < M < M_o$

(2) The left tail of the frequency curve is more elongated.

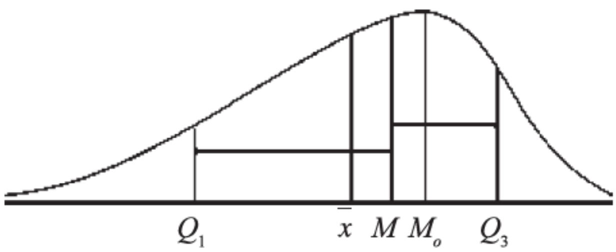
(3) $Q_3 - M = 18 - 15 = 3$ and $M - Q_1 = 15 - 10 = 5$. Hence $Q_3 - M < M - Q_1$

The following characteristics are generally seen in negatively skewed frequency distribution :

(1) Mean, median and mode are in increasing order. That is $\bar{x} < M < M_o$

(2) The distance between third quartile Q_3 and median M is less than the distance between median and first quartile Q_1. That is $(Q_3 - M) < (M - Q_1)$

(3) The left tail of the frequency curve of this distribution is more elongated.



Note: We find more variation in the observations to the left of the mode in a negatively skewed distribution. Many cases related to business and finance have negatively skewed frequency distribution. e.g. negatively skewed frequency curve is found in the frequency distributions of price of a certain item, number of investors, etc.

5.3 Concept of absolute and relative measure of skewness

We can find whether the skewness in the frequency distribution of population observations is positive or negative by drawing the frequency curve. But the degree of skewness in the distribution cannot be found using the graph.

Two types of measures are used to measure skewness: **(1) Absolute measure (2) Relative measure.** A measure of skewness which is expressed in the same units of population variable is called absolute measure of skewness. It is denoted by S_k.

The absolute measure of skewness is obtained by the difference of averages in Karl Pearson's method and by the difference of quartiles in Bowley's method. These measures cannot be used for comparing two populations having different units. Even when two populations have observations with same units, it is not advisable to use absolute measure as the measures of central tendency and dispersion may differ in the distributions of both the populations.

Thus, a relative measure is used for a comparative study of two or more populations which is called coefficient of skewness. The absolute measure of skewness of a population is divided by an appropriate measure of dispersion to obtain a measure of skewness which is free from the units. In short, **the relative measure of skewness is called coefficient of skewness.** It is denoted by j.

5.4 Methods for determining skewness and Coefficient of skewness

The following two methods are widely used to find skewness and coefficient of skewness of a frequency distribution: (1) Karl Pearson's method (2) Bowley's method.

5.4.1 Karl Pearson's method

The values of mean, median and mode are not same in a skewed distribution and the median is between mean and mode. Hence the difference between mean and mode is generally used to find the measure of skewness. Thus skewness $S_k = \bar{x} - M_o$. Coefficient of skewness j for a unimodal distribution is obtained by dividing skewness S_k by the standard deviation 's'. The coefficient of skewness is thus found by the following formula : $j = \frac{\bar{x} - M_o}{s}$

When a distribution has multiple modes or when the value of mode is ill-defined, the mode is obtained by empirical formula $M_o = 3M - 2\bar{x}$ given by Karl Pearson. Thus, skewness is obtained by the following formula:

Skewness S_k = mean – mode = $\bar{x} - M_o = \bar{x} - (3M - 2\bar{x}) = 3\bar{x} - 3M = 3(\bar{x} - M)$

and coefficient of skewness $j = \frac{3(\bar{x} - M)}{s}$.

Note : There are two types of skewness: (1) Positive skewness (2) Negative skewness

(1) If the mode is well-defined in a distribution, skewness is found by the formula $S_k = \bar{x} - M_o$. For positively skewed distribution we have $\bar{x} > M_o$ and thus $S_k = \bar{x} - M_o > 0$. Hence when $S_k > 0$, the distribution is said to have positive skewness. Similarly in case of ill-defined mode, $\bar{x} > M$, $S_k = 3(\bar{x} - M) > 0$ is considered as positive skewness.

(2) If the mode is well-defined in a distribution, skewness is found by the formula $S_k = \bar{x} - M_o$. For negatively skewed distribution we have $\bar{x} < M_o$ and thus $S_k = (\bar{x} - M_o) < 0$. Hence when $S_k < 0$, the distribution is said to have negative skewness. Similarly in case of ill-defined mode, $\bar{x} < M$, $S_k = 3(\bar{x} - M) < 0$ is considered as negative skewness.

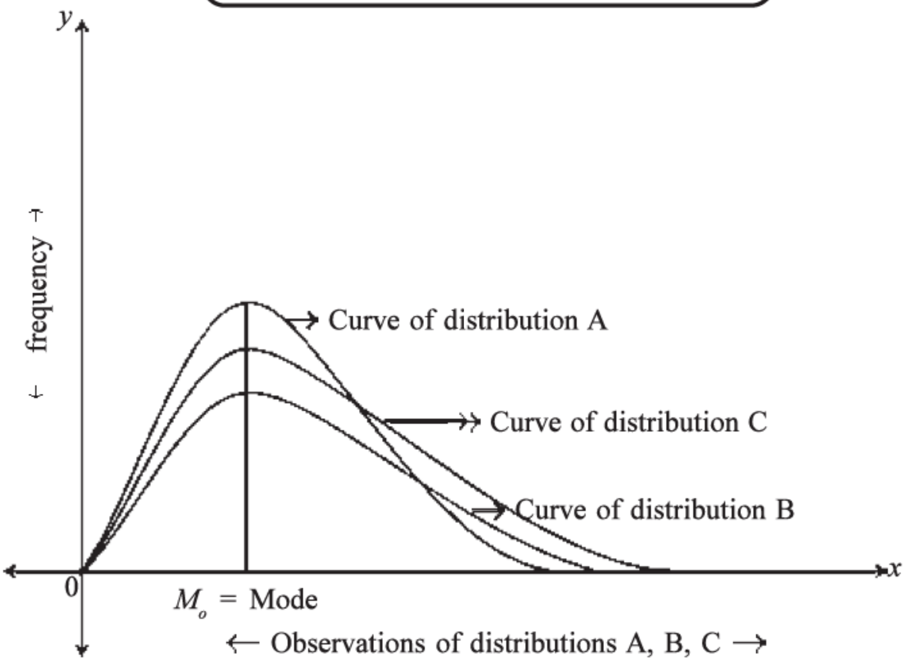
Note :

(1) In practice, coefficient of skewness lies between –1 and 1 for a skewed distribution based on sample data.

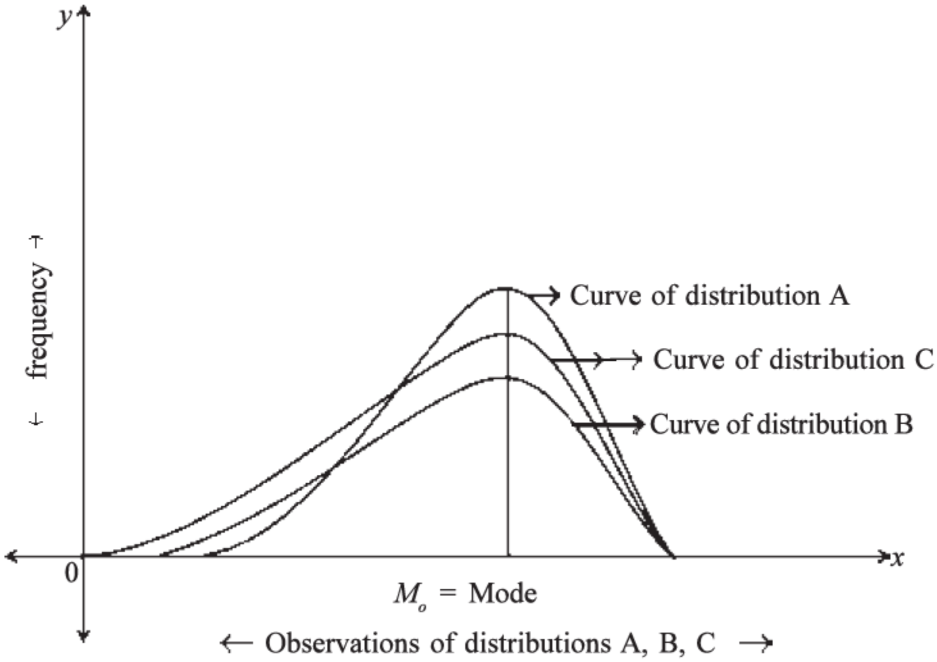
(2) If the frequency curve of a population has multiple modes, the coefficient of skewness $j = \frac{3(\bar{x} - M)}{s}$ lies between –3 and 3.

(3) In the year 1951 statistician N. L. Johnson proved that the coefficient of skewness obtained as $j = \frac{\bar{x} - M_o}{s}$ theoretically lies between $-\sqrt{3}$ and $\sqrt{3}$ that is between – 1.73 and 1.73 for a unimodal skewed distribution.

Additional Information for understanding



All the above frequency curves show positive skewness. The right tail for distribution C among them is the most elongated. So, it has maximum positive skewness. The curve of distribution B has less positive skewness than C. As the tail of distribution A is the least elongated as compared to B and C, it has the least positive skewness.



All the above frequency curves show negative skewness. The left tail for distribution C is the most elongated among them. So, it has the maximum negative skewness. The curve of distribution B has less negaitive skewness than C. As the tail of distribution A is the least elongated as compared to B and C, it has the least negaitive skewness.

Illustration 1 : The following data related to units transported by 50 trucks from railway yard to different factories on a day. Find the skewness and its coefficient using Karl Pearson's method from these data.

No. of units transported	120	130	140	150	160	170	180	190	200
No. of trucks	2	3	4	5	11	9	9	6	1

The given frequency distribution is unimodal. Hence, we will find mean, mode and standard deviation to find skewness by Karl Pearson's method.

No. of units transported x	**No. of trucks** f	$d = x - A$, $A = 160$	fd	fd^2
120	2	–40	–80	3200
130	3	–30	–90	2700
140	4	–20	–80	1600
150	5	–10	–50	500
160	11	0	0	0
170	9	10	90	900
180	9	20	180	3600
190	6	30	180	5400
200	1	40	40	1600
Total	$n = 50$		190	19500

Mean $\bar{x} = A + \frac{\Sigma fd}{n}$

$= 160 + \frac{190}{50}$

$= 160 + 3.8$

$= 163.8$

$\therefore \quad \bar{x} = 163.8$ units

Mode M_o = observation with the highest frequency = 160

$\therefore \quad M_o = 160$ units

Standard deviation $s = \sqrt{\frac{\Sigma fd^2}{n} - \left(\frac{\Sigma fd}{n}\right)^2}$

$= \sqrt{\frac{19500}{50} - \left(\frac{190}{50}\right)^2}$

$= \sqrt{390 - 14.44}$

$= \sqrt{375.56}$

$= 19.3794$

$\therefore \quad s \approx 19.38$ units

Skewness $S_k = \bar{x} - M_o$

$= 163.8 - 160$

$= 3.8$

$S_k = 3.8$ units

Coefficient of skewness $j = \frac{\bar{x} - M_o}{s}$

$= \frac{163.8 - 160}{19.38}$

$= \frac{3.8}{19.38}$

$= 0.1961$

$\therefore j \approx 0.20$

This distribution has positive skewness. It should be noted that $j = 0.20$ is free from units as coefficient of skewness is free from units.

Illustration 2 : Find the skewness and coefficient of skewness by Karl Pearson's method from the following data about annual tax of 100 companies.

Annual tax (lakh ₹)	0 – 20	20 – 40	40 – 60	60 – 80	80 – 100	100 – 120
No. of companies	5	20	40	25	6	4

As the given frequency distribution is unimodal, we will compute mean $\bar{x}$, mode M_o and standard deviation s.

Annual tax (lakh ₹)	No. of companies f	Mid-value x	$d = \frac{x - A}{c}$ $A = 50,\ c = 20$	fd	fd^2
0 – 20	5	10	– 2	– 10	20
20 – 40	20	30	– 1	– 20	20
40 – 60	40	50	0	0	0
60 – 80	25	70	1	25	25
80 – 100	6	90	2	12	24
100 – 120	4	110	3	12	36
Total	$n = 100$			**19**	**125**

Mean $\bar{x} = A + \frac{\Sigma fd}{n} \times c$

$= 50 + \frac{19}{100} \times 20$

$= 50 + \frac{380}{100}$

$= 50 + 3.8$

$\therefore \bar{x} = 53.8$

$\bar{x} =$ ₹ 53.8 lakh

Standard deviation $s = \sqrt{\frac{\Sigma fd^2}{n} - \left(\frac{\Sigma fd}{n}\right)^2} \times c$

$$= \sqrt{\frac{125}{100} - \left(\frac{19}{100}\right)^2} \times 20$$

$$= \sqrt{1.25 - 0.0361} \times 20$$

$$= \sqrt{1.2139} \times 20$$

$$= 22.0354$$

$$s \doteq ₹\ 22.04 \text{ lakh}$$

Mode M_o : The class interval with maximum frequency 40 is 40 - 60. The modal class in 40–60.

Now, $M_o = L + \frac{f_m - f_1}{2f_m - f_1 - f_2} \times c$

Here, $L = 40$, $f_m = 40$, $f_1 = 20$, $f_2 = 25$, $c = 20$

$$M_o = 40 + \frac{40-20}{2(40)-20-25} \times 20$$

$$= 40 + \frac{20 \times 20}{80-20-25}$$

$$= 40 + \frac{400}{35}$$

$$= 40 + 11.4285$$

$$= 51.4285$$

$$M_o \doteq ₹\ 51.43 \text{ lakh}$$

Skewness $S_k = \bar{x} - M_o$

$$= 53.8 - 51.43$$

$$= 2.37$$

$$S_k = ₹\ 2.37 \text{ lakh}$$

Coefficient of skewness $j = \frac{\bar{x} - M_o}{s} = \frac{53.8 - 51.43}{22.04}$

$$= \frac{2.37}{22.04}$$

$$= 0.1075$$

$$j \doteq 0.11$$

Hence, it can be said that the frequency distribution has positive skewness. Small value of coefficient of skewness indicates that it is almost symmetric.

Illustration 3 : A factory has 100 machines for production. The following data are obtained about rejected items during the production process. Find the skewness and its coefficient using Karl Pearson's method from these data.

No. of rejected items	11 – 20	21 – 30	31 – 40	41 – 50	51 – 60	61 – 70	71–80
No. of machines	2	12	25	39	12	9	1

The given frequency distribution is unimodal. Hence, we will compute mean, mode and standard deviation to find coefficient of skewness.

No. of rejected items	**No. of machines** f	**Mid-value** x	$d = \frac{x - A}{c}$ $A = 45.5$ $c = 10$	fd	fd^2
11 - 20	2	15.5	– 3	– 6	18
21 - 30	12	25.5	– 2	– 24	48
31 - 40	25	35.5	– 1	– 25	25
41 - 50	39	45.5	0	0	0
51 - 60	12	55.5	1	12	12
61 - 70	9	65.5	2	18	36
71 - 80	1	75.5	3	3	9
Total	$n = 100$			**– 22**	**148**

Mean $\bar{x} = A + \frac{\Sigma fd}{n} \times c$

$= 45.5 + \frac{(-22)}{100} \times 10$

$= 45.5 - 2.2$

$= 43.3$

$\therefore \quad \bar{x} = 43.3$ units

Mode M_o : The class interval with maximum frequency 39 is 41 - 50. Hence modal class of inclusive type is 41 - 50.

$\therefore$ Taking class boundaries, the modal class is 40.5 – 50.5.

Now, $M_o = L + \frac{f_m - f_1}{2f_m - f_1 - f_2} \times c$

Here, $L = 40.5, f_m = 39, f_1 = 25, f_2 = 12, c = 10$

$M_o = 40.5 + \frac{39-25}{2(39)-25-12} \times 10$

$= 40.5 + \frac{14 \times 10}{78 - 37}$

$= 40.5 + \frac{140}{41}$

$= 40.5 + 3.4146$

$= 43.9146$

$M_o \approx 43.91$ units

Standard deviation $s = \sqrt{\frac{\Sigma fd^2}{n} - \left(\frac{\Sigma fd}{n}\right)^2} \times c$

$= \sqrt{\frac{148}{100} - \left(\frac{-22}{100}\right)^2} \times 10$

$= \sqrt{1.48 - 0.0484} \times 10$

$= \sqrt{1.4316} \times 10$

$= 11.96$ units

Skewness $S_k = \bar{x} - M_o$

$= 43.3 - 43.91$

$S_k = -0.61$ units

Coefficient of skewness $j = \frac{\bar{x} - M_o}{s}$

$= \frac{43.3 - 43.91}{11.96}$

$= \frac{-0.61}{11.96}$

$= -0.0509$

$\therefore j \approx -0.05$

The given frequency distribution has negative skewness which is closer to symmetry.

Illustration 4 : The average monthly transportation cost of 200 families in a city in the year 2014 is as follows. Find skewness and coefficient of skewness using Karl Pearson's method.

Average monthly transportation cost (thousand ₹)	1 - 3	4 - 6	7 - 9	10 - 13	14 - 16	17 - 19
No. of families	5	40	120	20	10	5

As the given frequency distribution has classes of unequal lengths, we compute mean, median and standard deviation to find coefficient of skewness.

Average monthly transportation cost (thousand ₹)	No. of families f	Mid-value x	$d = x - A$, $A = 8$	fd	fd^2	cf
1–3	5	2	– 6	– 30	180	5
4–6	40	5	– 3	– 120	360	45
7–9	120	8	0	0	0	165
10–13	20	11.5	3.5	70	245	185
14–16	10	15	7	70	490	195
17–19	5	18	10	50	500	200
Total	$n = 200$			40	1775	

Mean $\bar{x} = A + \frac{\Sigma fd}{n}$

$= 8 + \frac{40}{200}$

$= 8 + 0.2$

$= 8.2$

$\therefore \quad \bar{x} =$ ₹ 8.2 thousand

Median M = value of the $\left(\frac{n}{2}\right)$ th observation

= value of the $\left(\frac{200}{2}\right)$th observation

= value of the 100 th observation

Referring to *cf* column, we find that the 100th observation lies in the interval 7 - 9. Hence, class of median is 7-9, which is inclusive type. ∴ The class boundaries of median class are 6.5 - 9.5

Now, $M = L + \frac{\frac{n}{2} - cf}{f} \times c$

Here, $L = 6.5$, $\frac{n}{2} = 100$, $cf = 45$, $f = 120$ and $c = 3$

$M = 6.5 + \frac{100 - 45}{120} \times 3$

$= 6.5 + \frac{55 \times 3}{120}$

$= 6.5 + 1.375$

$= 7.875$

$\therefore \quad M =$ ₹ 7.88 thousand

Standard deviation $s = \sqrt{\frac{\Sigma fd^2}{n} - \left(\frac{\Sigma fd}{n}\right)^2}$

$= \sqrt{\frac{1775}{200} - \left(\frac{40}{200}\right)^2}$

$= \sqrt{8.875 - 0.04}$

$= \sqrt{8.835}$

$= 2.9724$

$\therefore s \approx$ ₹ 2.97 thousand

Skewness $S_k = 3\ (\bar{x} - M)$

$= 3\ (8.2 - 7.88)$

$= 3\ (0.32)$

$= 0.96$

$S_k =$ ₹ 0.96 thousand

Coefficient of skewness $j = \frac{3(\bar{x} - M)}{s}$

$= \frac{3(8.2 - 7.88)}{2.97}$

$= \frac{3(0.32)}{2.97}$

$= \frac{0.96}{2.97}$

$= 0.3232$

$= 0.32$

$\therefore \quad j = 0.32$

Thus the given frequency distribution has positive skewness.

Illustration 5 : The following information is obtained for life (in complete hours) of 400 electric bulbs. Find the skewness and the coefficient of skewness by Karl Pearson's method.

Life of electric bulbs (completed hours)	4000-4199	4200-4399	4400-4599	4600-4799	4800-4999	5000-5199	5200-5399	5400-5599
No. of electric bulbs	14	46	58	76	70	76	40	20

Two classes of the above distribution have maximum frequency 76. Since this is a bimodal distribution, we will compute mean, median and standard deviation.

Life of electric bulbs (completed hours)	**No. of electric bulbs** f	**Mid-value** x	$d = \frac{x - A}{c}$, $A = 4699.5$, $c = 200$	fd	fd^2	cf
4000 - 4199	14	4099.5	– 3	– 42	126	14
4200 - 4399	46	4299.5	– 2	– 92	184	60
4400 - 4599	58	4499.5	– 1	– 58	58	118
4600 - 4799	76	4699.5	0	0	0	194
4800 - 4999	70	4899.5	1	70	70	264
5000 - 5199	76	5099.5	2	152	304	340
5200 - 5399	40	5299.5	3	120	360	380
5400 - 5599	20	5499.5	4	80	320	400
Total	$n = 400$			**230**	**1422**	

Mean $\bar{x} = A + \frac{\Sigma fd}{n} \times c$

$= 4699.5 + \frac{230}{400} \times 200$

$= 4699.5 + 115$

$= 4814.5$

$\therefore \quad \bar{x} = 4814.5$ hours

Median $M =$ value of the $\left(\frac{n}{2}\right)$ th observation

$=$ value of the $\left(\frac{400}{2}\right)$th observation

$=$ value of the 200th observation

Referring to *cf* column, we find that the 200th observation lies in the interval 4800 - 4999. Hence class of median is 4800 - 4999 which is inclusive type. ∴ The class boundaries of median class are 4799.5 - 4999.5.

Now, $M = L + \frac{\frac{n}{2} - cf}{f} \times c$

Here, $L = 4799.5$, $\frac{n}{2} = 200$, $cf = 194$, $f = 70$ and $c = 200$ છે.

$$M = 4799.5 + \frac{200-194}{70} \times 200$$

$$= 4799.5 + \frac{6 \times 200}{70}$$

$$= 4799.5 + \frac{120}{7}$$

$$= 4799.5 + 17.1429$$

$$= 4816.6429$$

$\therefore \quad M \fallingdotseq 4816.64$ hours

Standard deviation $s = \sqrt{\frac{\Sigma fd^2}{n} - \left(\frac{\Sigma fd}{n}\right)^2} \times c$

$$= \sqrt{\frac{1422}{400} - \left(\frac{230}{400}\right)^2} \times 200$$

$$= \sqrt{3.555 - 0.3306} \times 200$$

$$= \sqrt{3.2244} \times 200$$

$$= 359.1323$$

$\therefore \quad s = 359.13$ hours

Skewness $S_k = 3(\bar{x} - M)$

$$= 3(4814.5 - 4816.64)$$

$$= 3\ (-2.14)$$

$$= -\ 6.42$$

$\therefore \ S_k = -\ 6.42$ hours

Coefficient skewness $j = \frac{3(\bar{x} - M)}{s}$

$= \frac{3(4814.5 - 4816.64)}{359.13}$

$= \frac{-6.42}{359.13}$

$= -0.0178$

$j \fallingdotseq -0.02$

Thus, the given distribution has negative skewness.

Illustration 6 : The following information is about the annual advertisement cost (in lakh ₹) of 30 companies. Find skewness and coefficient of skewness using it.

Annual advertisement cost (lakh ₹)	0	3	5	8	10 - 20	20 - 30	30 - 40
No. of companies	3	4	5	10	5	2	1

The given distribution is mixed (discrete and continuous). Hence we find $\bar{x}$, M, s and coefficient of skewness.

Annual cost (lakh ₹)	**No. of companies f**	**Mid-value x**	**$d = x - A$ $A = 15$**	**fd**	**fd^2**	**cf**
0	3	0	– 15	– 45	675	3
3	4	3	– 12	– 48	576	7
5	5	5	– 10	– 50	500	12
8	10	8	– 7	– 70	490	22
10 - 20	5	15	0	0	0	27
20 - 30	2	25	10	20	200	29
30 - 40	1	35	20	20	400	30
Total	**$n = 30$**			**–173**	**2841**	

Mean $\bar{x} = A + \frac{\Sigma fd}{n}$

$= 15 + \frac{-173}{30}$

$= 15 - 5.7667$

$= 9.2333$

$\therefore \quad \bar{x} \approx$ ₹ 9.23 lakh

Median M = Value of the $\left(\frac{n}{2}\right)$th observation

= Value of the $\left(\frac{30}{2}\right)$th observation

= Value of the 15th observation

Referring to cf column, we find that the 15th observation is 8.

$\therefore M =$ ₹ 8 lakh

Standard deviation $s = \sqrt{\frac{\Sigma fd^2}{n} - \left(\frac{\Sigma fd}{n}\right)^2}$

$= \sqrt{\frac{2841}{30} - \left(\frac{-173}{30}\right)^2}$

$= \sqrt{94.7 - 33.2544}$

$= \sqrt{61.4456}$

$= 7.8387$

$\therefore \quad s \approx$ ₹ 7.84 lakh

Skewness $S_k = 3\,(\bar{x} - M)$

$= 3\,(9.23 - 8)$

$= 3(1.23)$

$= 3.69$

$\therefore \quad S_k =$ ₹ 3.69 lakh

Coefficient skewness $j = \frac{3(\bar{x} - M)}{s}$

$= \frac{3(9.23 - 8)}{7.84}$

$= \frac{3.69}{7.84}$

$= 0.4707$

$j \approx 0.47$

Thus, the given frequency distribution has positive skewness.

Illustration 7 : The frequency distribution of hourly wages paid to 600 workers is given below. Find the Karl Pearson's coefficient of skewness from this distribution.

Hourly wage (lakh ₹)	40 - 60	60 - 80	80 - 100	100 - 120	120 - 140	140 - 160	160 - 180	180 - 200	200 - 220	220 - 240
No. of workers (f)	10	12	16	20	50	60	72	100	120	140

The maximum frequency in this distribution is 140 which is for the last class instead of the central class. Hence, the mode is ill-defined. We will find the coefficient of skewness using the formula $j = \frac{3(\bar{x} - M)}{s}$

Hourly wages (₹)	No. of workers f	Mid-value x	$d = \frac{x - A}{c}$ $A = 130$ $c = 20$	fd	fd^2	cf
40 - 60	10	50	– 4	– 40	160	10
60 - 80	12	70	– 3	– 36	108	22
80 - 100	16	90	– 2	– 32	64	38
100 - 120	20	110	– 1	– 20	20	58
120 - 140	50	130	0	0	0	108
140 - 160	60	150	1	60	60	168
160 - 180	72	170	2	144	288	240
180 - 200	100	190	3	300	900	340
200 - 220	120	210	4	480	1920	460
220 - 240	140	230	5	700	3500	600
Total	$n = 600$			**1556**	**7020**	

Mean $\bar{x} = A + \frac{\Sigma fd}{n} \times c$

$= 130 + \frac{1556}{600} \times 20$

$= 130 + 51.8666$

$= 181.866$

$\therefore \bar{x} =$ ₹ 181.87

Standard deviation $s = \sqrt{\frac{\Sigma fd^2}{n} - \left(\frac{\Sigma fd}{n}\right)^2} \times c$

$= \sqrt{\frac{7020}{600} - \left(\frac{1556}{600}\right)^2} \times 20$

$= \sqrt{11.7 - 6.7254} \times 20$

$= \sqrt{4.9746} \times 20$

$= 44.6076$

$\therefore s =$ ₹ 44.61

Meadian M = Value of the $\left(\frac{n}{2}\right)$ th observation

= Value of the $\left(\frac{600}{2}\right)$th observation

= Value of the 300th observation

Referring to cf column, we find that the 300th observation lies in the interval 180-200. Hence, class of median is 180-200.

Now, $M = L + \frac{\frac{n}{2} - cf}{f} \times c$

Here, $L = 180$, $\frac{n}{2} = 300$, $cf = 240$, $f = 100$ and $c = 20$.

$$M = 180 + \frac{300 - 240}{100} \times 20$$

$$= 180 + \frac{60 \times 20}{100}$$

$$= 180 + 12$$

$$= 192$$

$\therefore$ $M =$ ₹ 192

Coefficient skewness $j = \frac{3(\bar{x} - M)}{s}$

$$= \frac{3(181.87 - 192)}{44.61}$$

$$= \frac{-30.39}{44.61}$$

$$= -0.6812$$

$$j \approx -0.68$$

Thus, the given frequency distribution has negative skewness.

Illustration 8 : The following figures are given to describe the changes in the market prices of shares of a company before and after their general body meeting. Take them into consideration to comment whether the proceedings of this meeting have affected the market pries of shares by computing coefficient of skewness.

Details	Before meeting	After meeting
No. of share transactions	6000	5800
Mean of prices of share (₹)	440	460
Median of prices of share (₹)	500	480
Standard deviation of prices of share (₹)	60	52

We will be able to comment on the changes in the distribution of market prices of shares before and after the general body meeting by computing the coefficient of skewness. As we are given mean and median, we will find the coefficient of skewness using the following formula.

Coefficient of skewness Before meeting :

$$j = \frac{3(\bar{x}-M)}{s}$$

$$= \frac{3(440-500)}{60}$$

$$= \frac{3(-60)}{60}$$

$$j = -3$$

Coefficient of skewness After meeting :

$$j = \frac{3(\bar{x}-M)}{s}$$

$$= \frac{3(460-480)}{52}$$

$$= \frac{3(-20)}{52}$$

$$j = -1.15$$

(1) The distributions of share prices in both situations have negative skewness.

(2) The coefficient of skewness has decreased after the general body meeting. Hence, we can say that there is a partial effect of general body meeting on the share prices.

Illustration 9 : The data about the number of runs scored by two cricketers in 10 matches are as follows. Use this information to determine which cricketer's game is more skewed.

Details	Cricketer A	Cricketer B
Mean $\bar{x}$	50	35
Mode M_o	56	31
Standard deviation s	14.4	5.2

We will compute the coefficient of skewness for the scores of both cricketers from the given data :

Coefficient of skewness for cricketer A :

$$j = \frac{\bar{x}-M_o}{s}$$

$$= \frac{50-56}{14.4}$$

$$= \frac{-6}{14.4}$$

$$j = -0.42$$

Coefficient of skewness for cricketer B :

$$j = \frac{\bar{x}-M_o}{s}$$

$$= \frac{35-31}{5.2}$$

$$= \frac{4}{5.2}$$

$$j = 0.77$$

Ignoring the signs of coefficient of skewness for scores of cricketers, we see that the value of coefficient for cricketer B is more than that of A. Hence, we say that the game of cricketer B is more skewed.

Illustration 10 : From the following measures obtained for the frequency distributions of sales (in lakh Rs.) of potatoes by two merchants in a month, determine which distribution is more close to symmetry.

Measures for distribution of sales of potatoes by merchant A	Measures for distribution of sales of potatoes by merchant B
Mean $\bar{x}$ = ₹ 40 Median M = ₹ 43 Standard deviation s = ₹ 25	Mean $\bar{x}$ = ₹ 45 Median M = ₹ 40 Standard deviation s = ₹ 16

Coefficient of skewness for sale of A :

$$j = \frac{3(\bar{x}-M)}{s} = \frac{3(40-43)}{25} = \frac{3(-3)}{25} = -0.36$$

$$j = -0.36$$

Coefficient of skewness for sale of B :

$$j = \frac{3(\bar{x}-M)}{s} = \frac{3(45-40)}{16} = \frac{3(5)}{16} = 0.9375$$

$$j = 0.94$$

The distribution having less numerical value (ignoring the sign) of coefficient of skewness is said to be more close to symmetry. The coefficient of distribution A is numerically less than that of B. Hence, the distribution of A is more close to symmetry than distribution B.

Illustration 11 : The median of the distribution of marks in statistics obtained by students of a school is 72 and its mean is 75. Find the skewness for the marks obtained by the students from these data and state the type of skewness.

Here, median M = 72 and mean $\bar{x}$ = 75. We will find the skewness using the following formula as the mode is not given.

Skewness $S_k = 3(\bar{x} - M)$
$= 3(75 - 72)$
$= 3(3) = 9$ Marks

$\therefore S_k = 9$ marks

As, $S_k > 0$ the frequency distribution of marks obtained by students has positive skewness.

Activity

Collect the data about the 'number of absent days in a month' for all the students of your class. Construct a frequency distribution for the number of absent days. Find the absolute and relative measures of skewness for it.

Find the absolute and relative measures of skewness for a similar distribution for the students of another class. Compare the frequency distributions of both the classes using frequency curve, coefficient of variation and coefficient of skewness.

EXERCISE 5.1

1. The following distribution shows demand of 500 ml pouches of pasteurized toned milk by 59 consumers. Find Karl Pearson's coefficient of skewness from these data.

Demand of milk pouches (units)	1	2	3	4	5	6	7	8
No. of consumers	2	7	10	15	12	7	4	2

2. The following distribution shows purchase of T-shirts by 270 customers according to the shoulder lengths (in inches). Find Karl Pearson's coefficient of skewness from these data and interpret it.

Shoulder length of shirts (inches)	12.0	12.5	13.0	13.5	14.0	14.5	15.0	15.5	16.0
No. of customers	5	20	30	47	56	56	37	16	3

3. The following information is obtained for the time taken (in completed minutes) by each worker to carry out a certain job. Find Karl Pearson's coefficient of skewness from these data and interpret it.

Time taken (completed min.)	5 - 9	10 - 14	15 - 19	20 - 24	25 - 29
No. of workers	3	8	4	2	1

4. The students of standard 11 have collected the data about profits (in crore Rs.) of IT companies. Find Karl Pearson's coefficient of skewness from these data and interpret it.

Profit (crore₹)	5 - 7	7 - 9	9 - 11	11 - 13	13 - 15	15 - 17
No. of companies	5	12	20	8	3	2

5. The following frequency distribution gives the amount of annual depreciation (in lakh ₹) of 38 companies. Using this information, find skewness and its coefficient by Karl Pearson's method. State the type of skewness.

Amount of annual depreciation (lakh ₹)	7	9	10	11 - 20	21 - 24	25 - 36
No. of companies	2	3	4	7	12	10

6. The monthly consumption of cotton (in thousand bales) of 35 cloth mills is as follows. Using this, find skewness and its coefficient by Karl Pearson's method. State the type of skewness.

Consumption of cotton (thousand bales)	0 - 2	2 - 6	6 - 12	12 - 20	20 - 22
No. of mills	3	10	7	12	3

7. The temperature (in Celsius) at a tourist place for 60 days in the year 2014 is as follows. Using this information, find skewness and its coefficient by Karl Pearson's method. State the type of skewness.

Temperature (Celsius)	–3° to –1°	–1° to 5°	5° to 11°	11° to 19°	19° to 23°	23° to 27°
No. of days	4	14	20	14	5	3

*

5.4.2 Bowley's Method

Prof. A. L. Bowley has given a measure of skewness based on quartiles. This measure is based on the position of quartiles. It considers the main assumption that 'both the quartiles Q_1 and Q_3 are not equidistant from the median $M (= Q_2)$ in a skewed frequency distribution. The absolute measure of skewness S_k is obtained using the quartile differences $Q_3 - M$ and $M - Q_1$.

Thus, skewness $= S_k = (Q_3 - M) - (M - Q_1)$

$\therefore \quad S_k = Q_3 + Q_1 - 2M$

We have studied two types of skewness. We know that if the frequency distribution has positive skewness $Q_3 - M > M - Q_1$.

$\therefore Q_3 + Q_1 > 2M$ and $S_k > 0$.

For a negatively skewed distribution, we have $Q_3 - M < M - Q_1$ $\therefore$ $Q_3 + Q_1 < 2M$ and $S_k < 0$.

If the frequency distribution is symmetric, we have $Q_3 - M = M - Q_1$ $\therefore$ $Q_3 + Q_1 = 2M$ and $S_k = 0$.

The distances of Q_1 and Q_3 from the median M are $M - Q_1$ and $Q_3 - M$ respectively. When the difference between these two values is divided by their sum, we obtain the relative measure of skewness called coefficient of skewness. Thus, Bowley's formula for **coefficient of skewness** j is as follows:

$$j = \frac{(Q_3 - M) - (M - Q_1)}{(Q_3 - M) + (M - Q_1)} = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$$

The measure of skewness obtained by this formula is called Bowley's coefficient of skewness. We know that the values of $(Q_3 - M)$ and $(M - Q_1)$ are positive. The difference between any two real numbers is less than or equal to their sum. Thus $\left|\frac{(Q_3 - M) - (M - Q_1)}{(Q_3 - M) + (M - Q_1)}\right| \le 1$.

$\therefore$ |Bowley's coefficient of skewness j | ≤ 1

$\therefore$ $-1 \le j \le 1$. Thus, the range of Bowley's coefficient of skewness is –1 to 1

Note :

(1) If the frequency distribution is open-ended, this is the only method to find absolute and relative measures of skewness.

(2) Bowley's method is appropriate when the skewness is to be found using the positional averages which are quartiles and median.

Illustration 12 : The data about rainfall (in cm) at a place in a month during monsoon are obtained as follows. Find the skewness and coefficient of skewness using Bowley's method.

Rainfall (cm)	6	7	13	5	15	20
No. of days	10	5	3	8	3	2

We will find quartiles Q_1, Q_2 (= M) and Q_3 after arranging the observations for rainfall in the ascending order and then compute skewness and its coefficient using Bowley's formula.

Rainfall (cm)	**No. of days** ***f***	***cf***
5	8	8
6	10	18
7	5	23
13	3	26
15	3	29
20	2	31
Total	$n = 31$	

First quartile Q_1 = Value of the $\left(\frac{n+1}{4}\right)$ th observation

= Value of the $\left(\frac{31+1}{4}\right)$th observation

= Value of the $\frac{32}{4}$th observation

= Value of the 8th observation

Referring to *cf* column, we find that the value of the 8th observation is 5.

$\therefore$ Q_1 = 5 cm

Median M = Value of the $\left(\frac{n+1}{2}\right)$ th observation

M = Value of the $\left(\frac{31+1}{2}\right)$ th observation

= Value of the $\frac{32}{2}$ th observation

= Value of the 16 th observation

Referring to *cf* column, we find that the value of the 16th observation is 6.

$\therefore$ M = 6 cm

Third quartile Q_3 = Value of the 3 $\left(\frac{n+1}{4}\right)$th observation

= Value of the 3 $\left(\frac{31+1}{4}\right)$th observation

= Value of the 3 $\left(\frac{32}{4}\right)$th observation

= Value of the 3(8)th observation

= Value of the 24th observation

Referring to *cf* column, we find that the value of the 24th observation is 13.

$\therefore$ Q_3 = 13 cm

Skewness S_k = $Q_3 + Q_1 - 2M$

= 13 + 5 – 2 (6)

= 18 – 12

= 6

S_k = 6 cm

Coefficient of skewness j = $\frac{Q_3+Q_1-2M}{Q_3-Q_1}$

= $\frac{13+5-2(6)}{13-5}$

= $\frac{18-12}{8}$

= $\frac{6}{8}$

j = 0.75

Illustration 13 : The frequency distribution of number of cheques received per day for clearing of 5 branches of a bank on 100 days in the year 2014 is as follows. Find the coefficient of skewness by Bowley's method using this distribution.

No. of cheques	0 - 199	200 - 399	400 - 599	600 - 799	800 - 999
No. of days	10	13	17	42	18

We will compute first quartile Q_1, median M (= Q_2) and third quartile Q_3 to find Bowley's coefficient.

No. of cheques	**No. of days** f	cf
0 - 199	10	10
200 - 399	13	23
400 - 599	17	40
600 - 799	42	82
800 - 999	18	100
Total	$n = 100$	

First quartile Q_1 = Value of the $\left(\frac{n}{4}\right)$th observation

= Value of the $\left(\frac{100}{4}\right)$th observation

= Value of the 25th observation

Referring to cf column, we find that the 25th observation lies in the interval 400 - 599. Hence class of Q_1 is 400 - 599 which is inclusive type. ∴ The class boundaries of class of Q_1 are 399.5 - 599.5

Now, $Q_1 = L + \frac{\frac{n}{4} - cf}{f} \times c$

Here, $L = 399.5$, $\frac{n}{4} = 25$, $cf = 23$ $f = 17$ and $c = 200$

$$Q_1 = 399.5 + \frac{25 - 23}{17} \times 200$$

$$= 399.5 + \frac{2 \times 200}{17}$$

$$= 399.5 + \frac{400}{17}$$

$$= 399.5 + 23.5294$$

$$= 423.0294$$

∴ Q_1 = 423.03 cheques

Median M = Value of the $\left(\frac{n}{2}\right)$th observation

= Value of the $\left(\frac{100}{2}\right)$th observation

= Value of the 50 th observation

Referring to cf column, we find that the 50th observation lies in the interval 600 - 799. Hence, class of median is 600-799 which is inclusive type. ∴ The class boundaries of class of Q_1 are 599.5 - 799.5.

Now, $M = L + \frac{\frac{n}{2} - cf}{f} \times c$

Here, $L = 599.5$, $\frac{n}{2} = 50$, $cf = 40$ $f = 42$ and $c = 200$

$$M = 599.5 + \frac{50 - 40}{42} \times 200$$

$$= 599.5 + \frac{10 \times 200}{42}$$

$$= 599.5 + \frac{2000}{42}$$

$$= 599.5 + 47.619$$

$$= 647.119$$

$$M \fallingdotseq 647.12 \text{ cheque}$$

Third quartile Q_3 = Value of the $3\left(\frac{n}{4}\right)$th observation

= Value of the $3\left(\frac{100}{4}\right)$th observation

= Value of the 3(25)th observation

= Value of the 75th observation

Referring to cf column, we find that the 75th observation lies in the interval 600 - 799. Hence, class of Q_3 is 600 - 799 which is inclusive type. ∴ The class boundaries of class of Q_3 are 599.5 - 799.5.

Now, $Q_3 = L + \frac{3\left(\frac{n}{4}\right) - cf}{f} \times c$

Here, $L = 599.5$, $3\left(\frac{n}{4}\right) = 75$, $cf = 40$, $f = 42$ and $c = 200$

$$Q_3 = 599.5 + \frac{75 - 40}{42} \times 200$$

$$= 599.5 + \frac{35 \times 200}{42}$$

$$= 599.5 + 166.6667$$

$$= 766.1667$$

$$\therefore \quad Q_3 \fallingdotseq 766.17 \text{ cheque}$$

Coefficient of skewness $j = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$

$$j = \frac{766.17 + 423.03 - 2(647.12)}{766.17 - 423.03}$$

$$= \frac{1189.20 - 1294.24}{343.14}$$

$$= \frac{-105.04}{343.14}$$

$$= -0.3061$$

$$j \fallingdotseq -0.31$$

Thus, we say that the frequency distribution has negative skewness.

Illustration 14 : Find the coefficient of skewness for the frequency distribution of sales of 500 companies in the year 2014-2015 using an appropriate method.

Sales (thousand tonnes)	Less than 4	4 - 7	7 - 10	10 - 13	13 - 16	16 - 20	20 and above
No. of companies	26	119	198	86	39	20	12

As the distribution has open-ended classes, measure of skewness can be obtained only by Bowley's method. We will compute first quartile, median and third quartile.

Sales (thousand tonnes)	No. of companies f	cf
Less than 4	26	26
4–7	119	145
7–10	198	343
10–13	86	429
13–16	39	468
16–20	20	488
20 and above	12	500
Total	$n = 500$	

First quartile Q_1 = Value of the $\left(\frac{n}{4}\right)$ th observation

= Value of the $\left(\frac{500}{4}\right)$th observation

= Value of the 125th observation

Referring to cf column, we find that the 125th observation lies in the interval 4 - 7. Hence, class of Q_1 is 4 - 7.

Now, $Q_1 = L + \frac{\frac{n}{4} - cf}{f} \times c$

Here, $L = 4$, $\frac{n}{4} = 125$, $cf = 26$, $f = 119$ and $c = 3$

$Q_1 = 4 + \frac{125 - 26}{119} \times 3$

$= 4 + \frac{297}{119}$

$= 4 + 2.4958$

$= 6.4958$

$\therefore \quad Q_1 \risingdotseq 6.50$ thousand tonnes

Median M = Value of the $\left(\frac{n}{2}\right)$th observation

= Value of the $\left(\frac{500}{2}\right)$th observation

= Value of the 250th observation

Referring to cf column, we find that the 250th observation lies in the interval 7 - 10. Hence class of median is 7 - 10.

$$\text{Now, } M = L + \frac{\frac{n}{2} - cf}{f} \times c$$

$$\text{Here, } L = 7, \ \frac{n}{2} = 250, \ cf = 145, \ f = 198 \text{ and } c = 3$$

$$M = 7 + \frac{250 - 145}{198} \times 3$$

$$= 7 + \frac{105 \times 3}{198}$$

$$= 7 + \frac{315}{198}$$

$$= 7 + 1.5909$$

$$= 8.5909$$

$$\therefore \quad M \doteq 8.59 \text{ thousand tonnes}$$

Third quartile Q_3 = Value of the $3\left(\frac{n}{4}\right)$ th observation

= Value of the $3\left(\frac{500}{4}\right)$th observation

= Value of the 3 (125)th observation

= Value of the 375th observation

Referring to cf column, we find that the 375th observation lies in the interval 10 - 13. Hence, class of Q_3 is 10 - 13.

$$\text{Now, } Q_3 = L + \frac{3\left(\frac{n}{4}\right) - cf}{f} \times c$$

$$\text{Here, } L = 10, \ 3\left(\frac{n}{4}\right) = 375, \ cf = 343, \ f = 86 \text{ and } c = 3$$

$$Q_3 = 10 + \frac{375 - 343}{86} \times 3$$

$$= 10 + \frac{32 \times 3}{86}$$

$$= 10 + \frac{96}{86}$$

$$= 10 + 1.1163$$

$$= 11.1163$$

$$\therefore \quad Q_3 \doteq 11.12 \text{ thousand tonnes}$$

Coefficient of skewness $j = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$

$$j = \frac{11.12 + 6.5 - 2(8.59)}{11.12 - 6.5}$$

$$= \frac{17.62 - 17.18}{4.62}$$

$$= \frac{0.44}{4.62}$$

$$= 0.0952$$

$$\therefore j \doteq 0.10$$

Thus, we say that the frequency distribution has positive skewness. We can say that the distribution is close to symmetry as the coefficient is close to zero.

Illustration 15 : For a frequency distribution of monthly overtime (in hours) of employees of a company, the difference between the two extreme quartiles is 50 and their sum is 218. If the median is 112, find the coefficient of skewness.

Here, $Q_3 + Q_1 = 218$, $Q_3 - Q_1 = 50$, $M = 112$

Coefficient of skewness $j = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$

$$= \frac{218 - 2(112)}{50}$$

$$= \frac{218 - 224}{50}$$

$$= \frac{-6}{50}$$

$$j = -0.12$$

Illustration 16 : Mode of a symmetric frequency distribution is 84. If the first quartile is 68, find the third quartile.

As the distribution is symmetric, we have $j = 0$ and $\bar{x} = M = M_o$

Thus, $M = M_o = 84$, $M = 84$ and $Q_1 = 68$

$$j = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$$

$$\therefore \quad 0 = \frac{Q_3 + 68 - 2(84)}{Q_3 - 68}$$

$$\therefore \quad 0\ (Q_3 - 68) = Q_3 + 68 - 168$$

$$\therefore \quad 0 = Q_3 - 100$$

$$\therefore \quad Q_3 = 100$$

Illustration 17 : The following information is available about marks in a subject obtained by the students of a school in the annual examination. 25% students have scored less than 28 marks whereas, 75% students have scored less than 47 marks. If the coefficient of skewness for marks is 0.4, find the median.

25 % observations have value less than 28 $\therefore Q_1 = 28$

75% observations have value less than 28 $\therefore Q_3 = 47, j = 0.4$

Now, $j = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$

$\therefore \quad 0.4 = \frac{47 + 28 - 2M}{47 - 28}$

$\therefore \quad 0.4 = \frac{75 - 2M}{19}$

$\therefore \quad 0.4 \times 19 = 75 - 2M$

$7.6 = 75 - 2M$

$\therefore \quad 2M = 75 - 7.6$

$\therefore \quad 2M = 67.4$

$\therefore \quad M = 33.7$ marks

5.5 Comparison of two methods of Coefiicient of Skewness

The values of coefficient of skewness computed by Karl Pearson's method and Bowley's method are generally not same as they are based on different averages. Mean, median and mode are used as averages along with standard deviation in Karl Pearson's method. On the other hand, positional averages namely quartiles are used in Bowley's method. Thus both the methods are based on different averages.

The calculation of coefficient of skewness by Bowley's method is easier than the calculation by Karl Pearson's method.

The value of coefficient of skewness obtained by Karl Pearson's method is more reliable than by Bowley's method as the coefficient by Bowley's method is based on quartiles. Quartiles are found using only central 50% observations of the data. Mean and standard deviation are used for the coefficient of skewness by Karl Pearson's method. These measures are based on all observations.

Skewness and its coefficient can only be found by Bowley's method for the frequency distribution with open-ended classes. Mean and standard deviation used in Karl Pearson's method cannot be found for a distribution with open-ended classes. ln Karl Pearson's method, the absolute measure $\bar{x} - M_o$ for skewness is divided by standard deviation to obtain the coefficient of skewness whereas, the absolute measure of skewness $(Q_3 - M) - (M - Q_1)$ is divided by $(Q_3 - Q_1)$ to find its coefficient by Bowley's method. We should also note that it is inappropriate to compare the values obtained by Karl Pearson's formula to the one computed by Bowley's formula. If the coefficient of skewness j =0, it indicates the absence of skewness which means that the distribution is symmetric. Although it is possible in some cases that the frequency curve of the distribution has positive or negative skewness even if $j = 0$.

Illustration 18 : The measures from the data about the monthly income of residents of two cities are obtained as follows. Find coefficient of skewness using Karl Pearson's and Bowley's method from these data.

Details	City A	City B
Mean $\bar{x}$	300	280
Median M	284	310
Standard deviation s	60	110
First quartile Q_1	124	160
Third quartile Q_3	390	520

City A :

Coefficient of skewness for Karl Pearson's method :

$$j = \frac{3(\bar{x}-M)}{s}$$

$$= \frac{3(300-284)}{60}$$

$$= \frac{3(16)}{60}$$

$$= \frac{48}{60}$$

$$j = 0.8$$

Bowley's method

$$j = \frac{Q_3+Q_1-2M}{Q_3-Q_1}$$

$$= \frac{390+124-2(284)}{390-124}$$

$$= \frac{514-568}{266}$$

$$= \frac{-54}{266}$$

$$= -0.203$$

$$j = -0.2$$

City B :

Coefficient of skewness for Karl Pearson's method :

$$j = \frac{3(\bar{x}-M)}{s}$$

$$= \frac{3(280-310)}{110}$$

$$= \frac{3(-30)}{110}$$

$$= \frac{-9}{11}$$

$$\therefore \ j \approx -0.82$$

Bowley's method

$$j = \frac{Q_3+Q_1-2M}{Q_3-Q_1}$$

$$= \frac{520+160-2\times 310}{520-160}$$

$$= \frac{680-620}{360}$$

$$= \frac{60}{360}$$

$$\therefore \ j \approx 0.17$$

The coefficient of skewness for city A is positive by Karl Pearson's method and negative by Bowley's method. On the other hand, it is negative by Karl Pearson's method and positive by Bowley's method for city B. Hence it is not appropriate to compare the coefficients obtained by two methods. It is advisable to compare using only one method.

EXERCISE 5.2

1. From the following frequency distribution of different youngsters excercising in a gymnasium, find coefficient of skewness by Bowely's method and state the type of skewness.

Age (years)	25	17	20	18	26	22	28	23
No. of youngsters	22	4	19	11	7	9	3	8

2. The frequency distribution of paid up share capital out of the issued share capital for 31 manufacturing companies is as follows. Find skewness and its coefficient by Bowley s method and state the type of skewness.

Paid up share capital (lakh ₹)	Less than 100	Less than 300	Less than 500	Less than 700	Less than 900	Less than 1100	Less than 1300
No. of companies	0	6	16	19	23	27	31

3. The distribution of sales (in thousand tonnes) of 400 companies during the year 2014 - 15 is as follows. Find skewness and its coefficient from these data and state the type of skewness.

Sales (thousand tonnes)	Less than 20	20–40	40–50	50–75	75–90	90–120	120 and above
No. of companies	30	70	125	100	40	20	15

4. The commission paid on insurance policy amount to agents in a branch of an insurance company during a month has the following frequency distribution. Find the coefficient of skewness by Bowley's method.

Commission paid (thousand Rs.)	10–12	12–14	14–16	16–18	18–20	20–22	22–24	24–26	26–28	28–30
No. of companies	4	10	16	29	52	80	32	23	17	1

*

Summary

- Generally there are two types of frequency distributions:
  (1) Symmetric distribution (2) Skewed distribution
- Values of mean, median and mode are equal in a symmetric distribution and generally it has bell-shaped frequency curve.
- The right or left tail of the frequency curve of a skewed distribution is more elongated.
- The frequency curve with left tail more elongated is called negatively skewed curve and the one with right tail more elongated is called positively skewed curve.
- Two types of measures are used for measuring skewness:
  (1) Absolute measure (skewness) (2) Relative measure (coefficient of skewness)
- Karl Pearson's method or Bowley's method is used for measuring skewness.
- Relative measure of skewness is called coefficient of skewness.
- The coefficient of skewness generally lies between – 1 and 1. In some specific cases, it lies between – 3 and 3.

List of formulae :

Formulae for Karl Pearson's method	Formulae for Bowley's method
Absolute measure (1) For well-defined mode Skewness $S_k = \bar{x} - M_o$ (2) For multiple modes or for ill-defined mode Coefficient of skewness $S_k = 3\ (\bar{x} - M)$	**Absolute measure** Skewness $S_k = (Q_3 - M) - (M - Q_1)$ $= Q_3 + Q_1 - 2M$
Relative measure (1) For well-defined mode Coefficient of skewness $j = \frac{\bar{x} - M_o}{s}$ (2) For multiple modes or for ill-defined mode Coefficient of skewness $j = \frac{3(\bar{x} - M)}{s}$.	**Relative measure** Coefficient of skewness $j = \frac{(Q_3 - M) - (M - Q_1)}{(Q_3 - M) + (M - Q_1)}$ $j = \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$

EXERCISE 5

Section A

Find the correct option for the following multiple choice questions:

1. Generally, what is the range of coefficient of skewness for data where the mode is ill-defined?

(a) 0 to 1 (b) –1 to +1 (c) –3 to +3 (d) 1 to 0

2. For a frequency distribution having negative skewness, what will be the value of its mean?

(a) More than mode (b) Less than mode

(c) Equal to mode (d) Nothing can be said about mean

3. The following measures are obtained for two distributions.

Distribution (i) Mean = 44, Median = 48 and Standard deviation = 20

Distribution (ii) Mean = 44, Median = 50 and Standard deviation = 24

Which of the following statements is true ?

(a) Distributions (i) and (ii) have same degree of skewness.

(b) Distribution (i) has more skewness than distribution (ii).

(c) Distribution (i) has less skewness than distribution (ii).

(d) Nothing can be said about skewness from the given data.

4. Two measures of central tendency are given for the following three frequency distributions. All of them are unimodal distributions. State the type of skewness for the three distributions.

(i) Distribution A: mode = 100 and mean = 116

(ii) Distribution B: median = 142.8 and mean = 142.8

(iii) Distribution C: median = 208 and mean = 192

(a) A is symmetric, B is negatively skewed and C is positively skewed.

(b) A is negatively skewed, B is positively skewed and C is symmetric.

(c) A is positively skewed, B is symmetric and C is negatively skewed.

(d) A is positively skewed, B is negatively skewed and C is symmetric.

5. Mode of a frequency distribution exceeds its mean by 2. What type of distribution is it ?

(a) Negatively skewed (b) symmetric (c) positively skewed (d) nothing can be said.

6. If $Q_3 + Q_1 = 60$ and $M = 30$ for a frequency distribution, which of the following statements about its skewness is true ?

(a) Distribution is highly skewed. (b) Distribution is less skewed.

(c) Distribution has lack of symmetry. (d) Distribution is symmetric.

7. In a moderately skewed distribution, (mean – mode) will be how many times (mean – median) ?

(a) 3 (b) –1 (c) $\frac{1}{3}$ (d) 0

8. Which of the following statements is false for a negatively skewed frequency distribution ?

(a) Value of mean is less than median and mode.

(b) The distance between the third quartile and median is less than the distance between median and the first quartile.

(c) The left tail of the frequency curve of the distribution is more elongated.

(d) The distance between the third quartile and median is more than the distance between median and the first quartile.

9. Which of the following statements is true for a symmetric distribution ?

(a) $Q_3 = 2M - Q_1$ (b) $Q_2 - Q_3 = Q_2 - Q_1$ (c) $Q_3 + Q_1 > 2M$ (d) $Q_3 + Q_1 < 2M$

10. If $(M - \bar{x}) = -\frac{1}{2}s$, find the value of j.

(a) $-\frac{1}{3}$ (b) $\frac{3}{2}$ (c) = –1.5 (d) 0.15

11. Which measure of central tendency has higher value in (i) negatively skewed distribution and (ii) positively skewed distribution?

(a) (i) mean (ii) mode (b) (i) median (ii) mode

(c) (i) mode (ii) mean (d) nothing can be said about mean, median and mode

12. If coefficients of skewness for distribution X is – 0.99 and that of distribution Y is 0.90. Which of the following statements is true ?

(a) Distribution X is more skewed. (b) Distribution Y is more skewed.

(c) Nothing can be said about skewness of X and Y. (d) Distributions X and Y have same skewness.

13. Which of the following statements is false ?

(a) If $Q_3 + Q_1 > 2M$ the distribution is positively skewed.

(b) Bowley's coefficient of skewness is found using positional averages.

(c) The absolute measure is divided by standard deviation in Karl Pearson's method to eliminate the effect of unit of the variable whereas in Bowley's method, the absolute measure is divided by the difference of quartiles.

(d) The frequencies of observations which are equidistant on both sides of the mode are equally distributed in a skewed distribution.

14. Which of the following statements is true ?

(a) The distribution in which observations on both sides of mode are equally distributed is called negatively skewed distribution.

(b) The distances of median from the extreme quartiles are same in a symmetric distribution.

(c) If $S_k > 0$, $\bar{x} > M$ and $\bar{x} < M_o$

(d) If $S_k < 0$, $\bar{x} < M$ and $\bar{x} > M_o$

Section B

Answer the following questions in one sentence :

1. What is meant by skewness ?
2. When do we call a frequency distribution to be symmetric ?
3. When do we say that a frequency distribution is skewed ?
4. What can you say about the position of median in a skewed frequency distribution ?
5. Explain how to determine the skewness using a frequency curve.
6. What is coefficient of skewness ? State the range for its value.
7. Which measures are used to obtain Karl Pearson's measure of skewness ?
8. State the basic assumption of obtaining Bowley's measure of skewness.
9. Which method gives more reliable value of coefficient of skewness ?
10. Coefficient of skewness is absolute measure or relative easure. Give reasons.
11. Which formula is used for finding coefficient of skewness when a distribution has open-ended classes ?
12. What is the range for coefficient of skewness computed using Karl Pearson's method for a frequency distribution with unequal class intervals ?
13. State the type of skewness for a frequency distribution whose three quartiles are 42, 36 and 40.
14. State the type of skewness for a frequency distribution where $(Q_3 - Q_2) < (Q_2 - Q_1)$.
15. The mean of a frequency distribution is less than its median by 2 units. State its type of skewness.
16. If $Q_3 + Q_1 = 125$ and $M = 62.5$ છે, for a set of data, what can be said about its skewness ?
17. If $\bar{x} = M = M_o = 48$ in a frequency distribution, what can you say about its coefficient ?

Section C

Answer the following questions :

1. Explain the types of skewness.
2. Show the positions of averages and quartiles using a diagram for each types of skewness.
3. State any two characteristics of symmetric distribution.
4. State any two characteristics of skewed distribution.
5. For a frequency distribution, skewness $S_k = -2.8$. If its mode is 48.8, find mean.
6. Sum of two extreme quartiles is 138 in a symmetric frequency distribution. Find its median.
7. The coefficient of a skewed frequency distribution is 0.75. If its standard deviation is is 20 and mean is 37.5, find median.
8. The mean of a skewed distribution exceeds its median by 3. If its coefficient of skewness is 0.75, find standard deviation.
9. If a frequency distribution has $Q_3 - Q_2 = 2\ (Q_2 - Q_1)$, find j.
10. If a frequency distribution has skewness $S_k = -6.6$ and quartiles deviation = 22, find j.
11. A skewed frequency distribution has mean = 40, mode = 46, $Q_3 + Q_1 = 76$ and $Q_3 - Q_1 = 20$. Find Bowley's coefficient of skewness.
12. 'The coefficient of skewness obtained by Bowley's method is not more reliable than the coefficient by Karl Pearson's method.' Explain this statement.
13. The three quartiles of a frequency distribution are 76, 98 and 40. Find j and state the type of skewness.
14. The coefficient of skewness of a distribution is 0.85. If its mean is 3.4 more than its mode, find its variance.
15. If a frequency distribution has $\bar{x} + M_o = 82$, $\bar{x} = 44$ and $s = 12$, find coefficient of skewness.

Section D

Answer the following questions :

1. Briefiy explain Karl Pearson's method of finding the coefficient of skewness.
2. Write a short note on Bowley's method of finding the coefficient of skewness.
3. State the circumstances in which Karl Pearson's formula $j = 3\left(\frac{\bar{x}-M}{s}\right)$ is used to find coefficient of skewness.
4. Differentiate between positive and negative skewness with details and using a diagram.
5. Which of the following populations is closer to symmetry ?

   Population A : $\bar{x} = 56$, $M_o = 60$ and $s = 24$

   Population B : $\bar{x} = 56$, $M = 60$ and $s = 30$
6. Find coefficient from the following data using an appropriate method and determine which population is more skewed among A and B.

   Population A : $4Q_1 = 3Q_2 = 2Q_3 = 144$

   Population B : $Q_1 = 34.8$, $Q_2 = 45.5$ and $Q_3 = 70$
7. Third quartile is at a distance of 12.8 from the median in a frequency distribution and its first quartile is at a distance of 11.2 from the median. Find skewness and its coefficient.

8. If coefficient of variation is 25 %, $\bar{x}$ = 32 and M_o = 32.2 for a set of data, find its coefficient of skewness.

9. Find coefficient of skewness from the following data :
$n = 20$, $\Sigma x = 640$, $\Sigma x^2 = 20,800$ and $M = 32.2$

10. Karl Pearson's coefficient of skewness for a data set is – 0.6. If mean = 60 and s = 10, find median and mode for the data.

11. Karl Pearson's skewness for a frequency distribution is 8 and coefficient of skewness is = $\frac{2}{3}$. If the mean is 64, find its median and coefficient of variation.

12. Find the coefficient of skewness for a frequency distribution with $Q_3 + Q_1 = 1.5\ M$ and $3\ (Q_3 - Q_1) = 2M$

13. A frequency distribution has $4\bar{x} = 6M_o = 144$, $s = 64$ and $Q_3 + Q_1 = 6\ (Q_3 - Q_1) = 60$. Find the coefficient using Karl Pearson's and Bowley's method.

Section E

Answer the following questions :

1. Define skewed frequency distribution and state its characteristics.

2. Define symmetric frequency distribution and state its characteristics.

3. Differentiate: Karl Pearson's and Bowley's method for coefficient of skewness.

4. Explain skewness and coefficient of skewness.

5. What are the main objectives of studying skewness ?

6. Various measures for the frequency distributions of monthly salaries of two production firms are as follows. Compare the firms using coefficient of skewness by Bowley's and Karl Pearson's method.

Details	Mean	Median	First quartile	Third quartile	Standard deviation
Firm A	350	344	324	356	26
Firm B	360	340	330	370	38

7. The frequency distribution of sale of notebooks from a stationary shop in the month of June for the year 2014 is as follows. Find coefficient of skewness using Karl Pearson's method.

Sale of notebooks (dozens)	30	25	21	20	18	16	15	12
No. of days	2	2	7	3	4	7	2	3

8. The following information is available about defective staplers after testing 50 packets of 500 staplers each. Find coefficient of skewness using Karl Pearson's method.

No. of defective staplers	19	20	21	22	23	24	25	26
No. of packets	5	18	10	8	4	2	2	1

9. For the frequency distribution of a set of data, $n = 200$, $\Sigma f\,(x - 240) = 0$, $\Sigma f\,(x - 240)^2 = 11,250$ and median = 246, find coefficient of skewness and state the type of skewness.

Section F

Solve the following :

1. The frequency distribution of number of accidents due to driving for more than the prescribed time is given below. Find coefficient of skewness by Bowley's method.

No. of driving hours more than prescribed time	4	3.5	3	2.5	2	1.5	1	0
No. of accidents	5	4	3	2	1	2	2	1

2. The daily temperature of a city in the year 2014 is recorded as follows. The daily temperature has not been below -10^0 C. Find Karl Pearson's coefficient of skewness. State the type of skewness.

Mid-value (Celsius)	–5	5	12	18	25
No. of days	25	35	105	125	75

3. The distribution of marks obtained by 60 students in an examination is as follows. Find coefficient of skewness by Karl Pearson's method and state the type of skewness.

Marks of students	0 - 10	10 - 20	20 - 30	30 - 40	40 - 50	50 - 60	Total
No. of students	5	12	38	38	20	7	120

4. The frequency distribution of profits earned by 150 companies during the year 2015-2016 is as follows. Find the coefficient of skewness using an appropriate method and state the type of skewness.

Profit (lakh ₹)	Less than 10	10 - 20	20 - 30	30 - 40	40 and above
No. of companies	15	30	50	40	15

5. The frequency distribution of demand of a certain item is as follows. Find skewness and coefficient of skewness by Karl Pearson's method.

Demand (units)	1	2	3	4 - 8	8 - 12	12 - 16	16 - 20
No. of customers	10	8	12	10	5	15	20

6. A sample of 50 screws was taken from the lots of screws produced at a factory to measure the diameter (in mm) of head of each screw and its frequency distribution is as follows. Find Bowley's coefficient of skewness and interpret it.

Diameter of head of screw (mm)	4 - 4.1	4 - 4.2	4 - 4.3	4 - 4.4	4 - 4.5	4–4.6	4–4.7	4–4.8
No. of screws	6	13	23	33	41	46	48	50

7. From the following frequency distribution of daily sales of packets of bread from a departmental store, find coefficient of skewness by Karl Pearson's method and interpret it.

No. of bread packets	0 - 3	3 - 5	5 - 10	10 - 15	15 - 20	20 - 30	30 - 40	40 - 60
No. of customers	15	12	8	6	4	3	2	1

8. The frequency distribution of sales from a shop selling glass is as follows. Find Bowley's coefficient of skewness and interpret it.

Size of glass (sq. m.)	1 - 1.9	2 - 2.9	3 - 3.9	4 - 4.9	5 - 5.9	6 - 6.9	7 - 7.9
No. of customers	10	40	20	50	30	30	20

9. A construction company builds houses with different areas. The frequency distribution of areas of houses is as follows. Find coefficient of skewness by Karl Pearson's method and interpret it.

Area of house (sq.m.)	100	140	180	220	260
No. of houses	10	25	50	25	10

10. The frequency distribution of units of power consumed in an hour for different machines during a production process of factory is as follows. Find coefficient of skewness by Bowley's method.

Units of power consumed	10 - 15	15 - 20	20 - 25	25 - 30	30 - 35	35 - 40
No. of machines	5	10	15	20	25	30

Arthur Lyon Bowley
(1869 - 1957)

Sir Arthur Lyon Bowley was a British Statistician and Mathematical Economist. Among the distinguished posts he held were those of Professor of Mathematics, Economics and Statistics at University College, Reading and London and Director of University of Oxford Institute of Statistics. His major works were "Three studies on the National Income", published in 1938 and "Wages and Income in the UK since 1860", published in 1937. He occupied, at various times, high positions in the Royal Statistical Society, The Royal Economic Society (elected fellow in 1893), the International Statistical Institute, the Econometric Society and the British Association for the Advancement of Science.

•