(*ii*) *Quartile coefficient of skewness* $= \dfrac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1}$

Its value always lies between – 1 and + 1.

(*iii*) *Coefficient of skewness based on third moment* $\gamma_1 = \sqrt{\beta_1}$.

where $\beta_1 = \mu_3^2/\mu_2^3$

Thus $\gamma_1 = \sqrt{\beta_1}$ gives the simplest measure of *skewness.*

25.11 KURTOSIS

Kurtosis measures the degree of peakedness of a distribution and is given by $\beta_2 = \mu_4/\mu_2^2$.

(*a*) *Leptokurtic* (*b*) *Platykurtic* (*c*) *Mesokurtic*

Fig. 25.5

$\gamma_2 = \beta_2 - 3$ gives the excess of *Kurtosis*. The curves with $\beta_2 > 3$ are called *Leptokurtic* and those with $\beta_2 < 3$ as *Platykurtic*. The normal curve for which $\beta_2 = 3$, is called *Mesokurtic* [Fig. 25.5].

Example 25.11. *The first four moments about the working mean 28.5 of a distribution are 0.294, 7.144, 42.409 and 454.98. Calculate the moments about the mean. Also evaluate β_1, β_2 and comment upon the skewness and kurtosis of the distribution.* (*V.T.U., 2[illegible] S*)

Solution. The first four moments about the arbitrary origin 28.5 are $\mu'_1 = 0.294$, $\mu'_2 = 7.144$, $\mu'_3 = 42.409$, $\mu'_4 = 454.98$.

$\therefore$
$$\mu'_1 = \frac{1}{N}\sum f_i(x_i - 28.5) = \frac{1}{N}\sum f_i x_i - 28.5 = \bar{x} - 28.5 = 0.294 \text{ or } \bar{x} = 28.794$$
$$\mu_2 = \mu'_2 - \mu'^2_1 = 7.144 - (0.294)^2 = 7.058$$
$$\mu_3 = \mu'_3 - 3\mu'_2\mu'_1 + 2\mu'^3_1 = 42.409 - 3(7.144)(0.294) + 2(0.294)^3 = 36.151.$$
$$\mu_4 = \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2\mu'^2_1 - 3\mu'^4_1$$
$$= 454.98 - 4(42.409) \times (0.294) + 6(7.144)(0.294)^2 - 3(0.294)^4 = 408.738$$

Now
$$\beta_1 = \mu_3^2/\mu_2^3 = (36.151)^2/(7.058)^3 = 3.717$$
$$\beta_2 = \mu_4/\mu_2^2 = 408.738/(7.058)^2 = 8.205.$$

$\therefore$ $\gamma_1 = \sqrt{\beta_1} = 1.928$, which indicates considerable skewness of the distribution.

$\gamma_2 = \beta_2 - 3 = 5.205$ which shows that the distribution is *leptokurtic.*

Example 25.12. *Calculate the median, quartiles and the quartile coefficient of skewness from the following data :*

Weight (lbs) :	*70–80*	*80–90*	*90–100*	*100–110*	*110–120*	*120–130*	*130–140*	*140–150*
No. of persons :	*12*	*18*	*35*	*42*	*50*	*45*	*20*	*8*

Solution. Here total frequency $N = \Sigma f_i = 230$.

The cummulative frequency table is

Weight (lbs) :	*70–80*	*80–90*	*90–100*	*100–110*	*110–120*	*120–130*	*130–140*	*140–150*
f :	12	18	35	42	50	45	20	8
cum. f. :	12	30	65	107	157	202	222	230

Now $N/2 = 230/2 = 115$th item which lies in 110–120 group.

$\therefore$
$$\text{median or } Q_2 = L + \frac{N/2 - C}{f} \times h = 110 + \frac{115 - 107}{50} \times 10 = 111.6$$

Also $N/4 = 230/4 = 57.5$ *i.e.* Q_1 is 57.5th or 58th item which lies in 90–100 group.

$$\therefore \qquad Q_1 = L + \frac{N/4 - C}{f} \times h = 90 + \frac{57.5 - 30}{35} \times 10 = 97.85$$

Similarly, $3N/4 = 172.5$ *i.e.* Q_3 is 173rd item which lies in 120–130 group.

$$\therefore \qquad Q_3 = L + \frac{3N/4 - C}{f} \times h = 120 + \frac{172.5 - 157}{45} \times 10 = 123.44$$

Hence quartile coefficient of skewness $= \dfrac{Q_1 + Q_3 - 2Q_2}{Q_3 - Q_1}$

$$= \frac{97.85 + 123.44 - 2 \times 111.6}{123.44 - 97.85} = -0.07 \text{ (approx.).}$$

PROBLEMS 25.3

1. Calculate the first four moments of the following distribution about the mean :

x :	0	1	2	3	4	5	6	7	8
f:	1	8	28	56	70	56	28	8	1

Also evaluate β_1 and β_2. *(V.T.U., 2004 ; Madras, 2003)*

2. The following table gives the monthly wages of 72 workers in a factory. Compute the standard deviation, quartile deviation, coefficients of variation and skewness. *(V.T.U., 2001)*

Monthly wages (in ₹)	*No. of workers*	*Monthly wages (in ₹)*	*No. of workers*
12.5—17.5	2	37.5—42.5	4
17.5—22.5	22	42.5—47.5	6
22.5—27.5	19	47.5—52.5	1
27.5—32.5	14	52.5—57.5	1
32.5—37.5	3		

3. Find Pearson's coefficient of skewness for the following data :

Class :	10–19	20–29	30–39	40–49	50–59	60–69	70–79	80–89
Frequency :	5	9	14	20	25	15	8	4

(V.T.U., 2000 S)

4. Compute the quartile coefficient of skewness for the following distribution :

x :	3–7	8–12	13–17	18–22	23–27	28–32	33–37	38–42
f:	2	108	580	175	80	32	18	5

(Madras, 2002 ; V.T.U., 2000)

Also compute the measure of skewness based on the third moment.

5. The first three moments of a distribution about the value 2 of the variable are 1, 16 and – 40. Show that the mean = 3, the variance = 15 and $\mu_3 = -86$. *(V.T.U., 2003 S)*

6. Compute skewness and kurtosis, if the first four moments of a frequency distribution $f(x)$ about the value $x = 4$ are respectively 1, 4, 10 and 45. *(Coimbatore, 1999)*

7. In a certain distribution, the first four moments about a point are – 1.5, 17, – 30 and 108. Calculate the moments about the mean, β_1 and β_2 ; and state whether the distribution is leptokurtic or platykurtic ?

25.12 CORRELATION

So far we have confined our attention to the analysis of observations on a single variable. There are, however, many phenomenae where the changes in one variable are related to the changes in the other variable. For instance, the yield of a crop varies with the amount of rainfall, the price of a commodity increases with the reduction in its supply and so on. Such a simultaneous variation, *i.e.* when the changes in one variable are associated or followed by changes in the other, is called *correlation*. Such a data connecting two variables is called *bivariate population.*

If an increase (or decrease) in the values of one variable corresponds to an increase (or decrease) in the other, the correlation is said to be *positive*. If the increase (or decrease) in one corresponds to the decrease (or increase) in the other, the correlation is said to be *negative*. If there is no relationship indicated between the variables, they are said to be *independent or uncorrelated.*