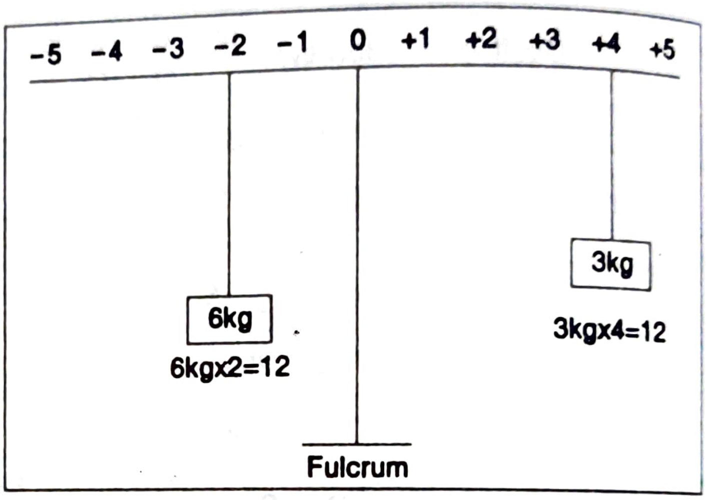
MOMENTS

'Moment' is a familiar mechanical term which refers to the measure of a force with respect to its tendency to provide rotation. The strength of the tendency depends on the amount of force and the distance from the origin of the point at which the force is exerted. If a number of forces, F_1, F_2, F_n at distances X_1, X_2, X_n are applied, the moment of the first force about the origin is $F_1 X_1$, the moment of the second force is $F_2 X_2$ etc. These moments are additive so that ΣFX is the total moment about the origin. If the total moment is divided by the total force, the quotient is termed "a moment". The formula is $\frac{\Sigma FX}{N}$ where $N = \Sigma F$ is the total force.

The following diagram shows this fact.



If on both the sides of the fulcrum the forces (weight × distance) are equal, then there will be a balance. At a balanced position the positive product equals the negative product.

However, the term moment as used in physics has nothing to do with the moment used in statistics, the only analogy being that in statistics we talk of moment of random variable about some point. The moments in statistics are used to describe the various characteristics of a frequency distribution like central tendency, variation, skewness and kurtosis.

It can be seen that the formula for a moment coefficient is identical with that for an arithmetic mean. This identity has led statisticians to speak of the arithmetic mean as the "*first moment about the origin*". Technically the mean is moment coefficient and not a total moment, but in the case of frequency curves, with which mathematical statistics is primarily concerned, the total frequency N is generally taken as unity, so the total moment and the moment coefficient are identical. In any case it has become customary in statistics to speak of the mean $\overline{X} = \frac{\Sigma FX}{N}$ as the first moment about the origin, and distinction between the total moment and moment coefficient is ignored. The concept of moments is also extended to higher powers. The statistical definition of the term 'moment' is given below :

Let the symbol x be used to represent the deviation of any item in a distribution from the arithmetic mean of that distribution. The arithmetic mean of the various powers of these deviations in any distribution is called the moment of the distribution. If we take the mean of the first power of the deviations we get the first moment about the mean : the mean of the squares of the deviations gives us the second moment about the mean; the mean of the cubes of the deviations gives us the third moment about the mean; and so on. The moments about mean called the 'central moment' are denoted by Greek letter μ (read as *mu*): thus μ_1 stands for first moment about mean. μ_2 stands for second moment about mean, etc. Symbolically,

$$\mu_1 = \frac{\Sigma(X - \bar{X})}{N} \text{ or } \frac{\Sigma x}{N}$$ [since sum of deviation of items from arithmatic mean is always zero, μ_1 would always be zero]

$$\mu_2 = \frac{\Sigma(X - \bar{X})^2}{N} \text{ or } \frac{\Sigma x^2}{N}$$ [$\mu_2 = \sigma^2$ = or $\sigma = \sqrt{\mu_2}$]

$$\mu_3 = \frac{\Sigma(X - \bar{X})^3}{N} \text{ or } \frac{\Sigma x^3}{N}$$

For frequency distribution

$$\mu_1 = \frac{\Sigma f(X - \bar{X})}{N} \text{ or } \frac{\Sigma fx}{N}; \quad \mu_2 = \frac{\Sigma f(X - \bar{X})^2}{N} \text{ or } \frac{\Sigma fx^2}{N}$$

$$\mu_3 = \frac{\Sigma(X - \bar{X})^3}{N} \text{ or } \frac{\Sigma fx^3}{N}; \quad \mu_4 = \frac{\Sigma f(X - \bar{X})^4}{N} \text{ or } \frac{\Sigma fx^4}{N}$$

Moments can be extended to higher powers in a similar way, but generally in practice the first four moments suffice. Furthermore, as pointed out by Yule and Kendall, "moment of higher order, though important in theory, are so extremely sensitive to sampling fluctuations that values calculated for moderate number of observations are quite unreliable and hardly ever repay the labour of computation".

Two important constants of a distribution are calculated from μ_2, μ_3 and μ_4. They are :

(*i*) β_1 (read as beta one) $= \dfrac{\mu_3^2}{\mu_2^3}$

(*ii*) β_2 (read as beta two) $= \dfrac{\mu_4}{\mu_2^2}$

β_1 measures skewness and β_2 kurtosis.

In a symmetrical distribution all odd moments, *i.e.*, μ_1, μ_3 etc. would always be zero. The reason is that if the curve is symmetrical there will be a deviation below the mean which exectly equals each deviations above the mean and, therefore, positive deviations and negative deviation will exactly balance out and when added will cancel out, *i.e.*, $\Sigma(X - \bar{X})$ would always be zero. Of course, if the deviations are raised to even powers, their sign will always be positive and they will no longer cancel out. But the sum of the odd powers will all be equal to zero on account of the cancellations. Thus, *odd moments are always zero in symmetrical distribution*. However, this rule does not hold true in asymmetrical distribution.

Moments about Arbitrary Origin

Where the actual mean is in fractions it is difficult to calculate moments by applying the above formulae. In such a case we can first compute moments about an arbitrary origin '*A*' and then convert these moments into moments about the actual mean. Moments about arbitrary origin, also called 'raw moments', are denoted by the symbol μ' to distinguish them from the moments about mean which are denoted by μ. Thus μ_1' would stand for the first moment about an arbitrary '*A*', μ_2' for the second moment about an arbitrary origin A, and so on. The calculations shall be done as follows :

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

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

For a frequency distribution

$$\mu_1' = \frac{\Sigma(X-A)}{N} \text{ or } \frac{\Sigma fd}{N} \text{ or } \frac{\Sigma fd}{N}\times i \qquad \left[\text{where } d = \frac{(X-A)}{N}\right]$$

$$\mu_2' = \frac{\Sigma f(X-A)^2}{N} \text{ or } \frac{\Sigma fd^2}{N} \text{ or } \frac{\Sigma fd^2}{N}\times i^2$$

In the same way the formulae for 3rd and 4th moments can be written.

Conversion of Moments about an Arbitrary Origin into Moments about Mean or Central Moments

For the sake of simplicity in calculations moments are first calculated about an arbitrary origin. If we want to obtain moments about mean we can do so with the following relationship :

$$\mu_1 = \mu_1' - \mu_1' = 0;\ \mu_3 = \mu_3' - 3\mu_1'\,\mu_2' + 2\,(\mu_1')^3$$

$$\mu_2 = \mu_2' - (\mu_1')^2\,;\ \mu_4 = \mu_4' - 4\mu_1'\,\mu_3' + 6\,(\mu_1')^2\,(\mu_2') - 3\,(\mu_1')^4$$

Relationship between the moments about mean and in terms of moments about any arbitrary point and conversely

$$\mu_r = \frac{1}{N}\Sigma(X_i-\bar{X})^r,\quad \mu_i' = \frac{1}{N}\Sigma(X_i-a)^2$$

$$X_i - \bar{X} = (X_i - a) - (\bar{X} - a)$$

$$\mu_r = \frac{1}{N}\Sigma(X_i-d)^r$$

where $X_i = x_i - a$ and $d = \bar{X} - a$

Using Binomial theorem

$$= \frac{1}{N}[\Sigma X_i^r - {}^rC_1 d\,\Sigma X_i^{r-1} + {}^rC_2 d^2\,\Sigma X_i^{r-2} \ldots + {}^rC_{r-1}(-d)^{r-1}\,\Sigma X_i + (-d)^r]$$

$$= \mu_r^1 - {}^rC_1\,d\mu'_{r-1} + {}^rC_2 d^3\mu'_{r-2} + \ldots + (-1)^{r-1}\,{}_r d^{r-1}\,\mu_1' + (-1)^r\,d^r$$

$$\mu_r = \mu_r' - {}^rC_1\,\mu_1'{}_{r-1} + {}^rC_2\,\mu_1'^2\,\mu'_{r-2} + (-1)^{r-1}\,(r-1)\,\mu_1'^r$$

Putting $r = 1, 2, 3, 4$ we get

$$\mu_1 = \mu_1' - \mu_1' = 0$$

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

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

$$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6\,(\mu_1')^2\,\mu_2' - 3\,(\mu_1')^4$$

Conversely,

$$\mu_r' = \frac{1}{N}\Sigma(X_i-a)^r$$

$$= \frac{1}{N}\Sigma(X-\bar{X}+\bar{x}-a)^r$$

$$\frac{1}{N}\Sigma(x_1' - d)^r \text{ where } x_i' = x_i - \bar{x} \text{ and } d = \bar{x} - a$$

$$= \frac{1}{N}[\Sigma x_1'^r + {}^rC_1\, d\Sigma x_i'^{r-1} + {}^rC_2\, d^2 \Sigma x_i'^{r-2} + \ldots + {}^rC_{r-1} a^{r-1} \Sigma x_1' + a^r]$$

$$= \mu_r + {}^rC_1 \mu_r - 1 + d^r C_2 \mu_{r-2} d^2 + \ldots + {}^rC_2 \mu_2 d^{r-2} + d^r$$

In particular

$$\mu_2' = \mu_2 + d^2$$

$$\mu_3' = \mu_3' - 3d\mu_2 + d^3$$

$$\mu_4' = \mu_1 + 4d\mu_3 + 6d^2 \mu_2 + d^4$$

Moments about Zero

The moments about zero are often denoted by $\upsilon_1, \upsilon_2, \upsilon_3$ and are obained as follows :

$$\upsilon_1 = \frac{\Sigma f X}{N} \qquad \upsilon_2 = \frac{\Sigma f X^2}{N}$$

$$\upsilon_3 = \frac{\Sigma f X^3}{N} \qquad \upsilon_4 = \frac{\Sigma f X^4}{N}$$

Also :

The first moment about zero or $\upsilon_1 = A + \mu_1'$ or the mean*

The second moment about zero or $\upsilon_2 = \mu_2 + (\upsilon_1)^2$.

The third moment about zero or $\upsilon_3 = \mu_3 + 3\upsilon_1\upsilon_2 - 2\upsilon_1^3$

The fourth moment about zero or $\upsilon_4 = \mu_4 + 4\upsilon_1\upsilon_3 - 6\upsilon_1^2\upsilon_2 + 3\upsilon_1^4$

Purpose of Moments. The concept of moment is of great significance in statistical work. With the help of moments we can measure the central tendency of a set of observations, their variability, their asymmetry and the height of the peak their curve would make. Because of the great convenience in obtaining measures of the various characteristics of a frequency distribution, the calculation of the first four moments about the mean may well be made the first step in the analysis of a frequency distribution.

The following is the summary of how moments help in analysing a frequency distribution :

Moment	*What it measures*
1. First moment about origin	Mean
2. Second moment about the mean	Variance
3. Third moment about the mean	Skewness
4. Fourth moment about the mean	Kurtosis

Sheppard's Correction for Grouping Errors

While calculating moments it is assumed that all the values of a variable in a class interval are concentrated at the centre of that interval (*i.e.*, mid-point). However, in practice, it is not so—the assumption is an approximation to facilitate calculations and it introduces some error which is known as *grouping error*. But for distributions of symmetrical or moderate-

* The signs are reverse of what we had while converting moments about an arbitrary origin to moments about mean.

ly skew type and class intervals not greater than about one-twentieth of the range, the approximation may be very close to one. In other cases we should apply Sheppard's corrections to eliminate grouping error.

The moments that we have computed, which have not been corrected by Sheppard's process, are called the *crude moments* to distinguish them from the *adjusted moments* which we get by applying Sheppard's corrections. These corrections are :

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

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

where i is the width of the class interval.

The first and third moments need no correction.

Conditions for applying Sheppard's Corrections

The following conditions should be satisfied for the application of Sheppard's corrections :

- The correction should not be made unless the frequency is at least 1,000 otherwise the moments will be more affected by sampling errors than by grouping errors.
- The correction is not applicable to J- or U-shaped distributions or even to the skew form.
- The observations should relate to a continuous variable.
- The frequencies should taper off to zero in both directions, *i.e.*, the curve should approach the base line gradually and slowly at each end of the distribution.

However, as pointed out by A.E. Waugh, the corrections are small and the statistician is *foolish* to bother with them if the original figures are rough approximations. But where we have continuous data with the characteristics described above and where the original measurements are reasonaly precise, we may well apply Sheppard's corrections to eliminate the grouping error.

Illustration 8. Analyse the frequency distribution by the method of moments :

X:	2	3	4	5	6
f:	1	3	7	3	1

Solution. CALCULATION OF FIRST FOUR MOMENTS

X	f	$(X-\bar{X})$ $\bar{X}=4$ x	fx	fx^2	fx^3	fx^4
2	1	−2	−2	4	−8	16
3	3	−1	−3	3	−3	3
4	7	0	0	0	0	0
5	3	+1	+3	3	+3	3
6	1	+2	+2	4	+8	16
	$N = 15$		$\Sigma fx = 0$	$\Sigma fx^2 = 14$	$\Sigma fx^3 = 0$	$\Sigma fx^4 = 38$

$$\mu_1 = \frac{\Sigma(X-\bar{X})}{N} \text{ or } \frac{\Sigma fx}{N} = \frac{0}{15} = 0; \quad \mu_3 = \frac{\Sigma fx^3}{N} = 0$$

$$\mu_2 = \frac{\Sigma fx^2}{N} = \frac{14}{15} = 0.933; \qquad \mu_4 = \frac{\Sigma fx^4}{N} = \frac{38}{15} = 2.533$$

$$\sigma = \sqrt{\text{variance}} \text{ or } \sqrt{\mu_2}; \qquad \sqrt{\mu_2} = \sqrt{0.933} = 0.966$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{0}{(0.933)^2} = 0$$

In a symmetrical distribution β_1 is zero. Hence, this distribution is symmetrical,

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{2.532}{(0.933)^2} = \frac{2.533}{0.87} = 2.91$$

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

Illustration 9. Calculate first four moments about the mean and also the value of β_1 and β_2 from the following data :

Marks :	0—10	10—20	20—30	30—40	40—50	50—60	60—70
No. of students :	8	12	20	30	15	10	5

(MBA, Kumaun Univ., 2001)

Solution. CALCULATION OF FIRST FOUR MOMENTS AND β_1 AND β_2

Marks	*m.p.* *m*	*No. of students* *f*	*(m – 35) / 10* *d*	*fd*	fd^2	fd^3	fd^4
0–10	5	8	–3	–24	72	–216	648
10–20	15	12	–2	–24	48	–96	192
20–30	25	20	–1	–20	20	–20	20
30–40	**35**	30	0	0	0	0	0
40–50	45	15	+1	+15	15	+15	15
50–60	55	10	+2	+20	40	+80	160
60–70	65	5	+3	+15	45	+135	405
		$N = 100$		$\Sigma fd = -18$	$\Sigma fd^2 = 240$	$\Sigma fd^3 = -102$	$\Sigma fd^4 = 1440$

$$\mu_1' = \frac{\Sigma fd}{N} \times i = \frac{-18}{100} \times 10 = -1.8$$

$$\mu_2' = \frac{\Sigma fd^2}{N} \times i^2 = \frac{240}{100} \times 100 = 240$$

$$\mu_3' = \frac{\Sigma fd^3}{N} \times i^3 = \frac{-102}{100} \times 1000 = -1020$$

$$\mu_4' = \frac{\Sigma fd^4}{N} \times i^4 = \frac{1440}{100} \times 10000 = 144000$$

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

$$\mu_2 = \mu_2' - (\mu_1')^2 = 240 - (-1.8)^2 = 240 - 3.24 = 236.76$$

$$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2\mu_1'^3 = -1020 - 3(1.8)(240) + 2(-1.8)^3$$
$$= -1020 + 1296 - 11.664 = 264.336$$

$$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6(\mu_1')^2\mu_2' - 3\mu_1'^4$$
$$= 144000 - 4(-1.8)(-1020) + 6(-1.8)^2(240) - 3(-1.8)^4$$
$$= 144000 - 7344 + 4665.6 - 31.4928 = 141290.11$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(264.336)^2}{(236.76)^3} = 0.005; \quad \beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{141290.11}{(236.76)^2} = 2.521.$$

Illustration 10. From the following data calculate moments about :
(*i*) assumed mean 25, (*ii*) actual mean, (*iii*) moments about zero.

Variable :	0—10	10—20	20—30	30—40
Frequency :	1	3	4	2

(B. Com., M.G. Univ., 1997)

Solution. CALCULATION OF MOMENTS ABOUT ARBITRARY ORIGIN

Variable	m	f	$(m-25)/10$ d	fd	fd^2	fd^3	fd^4
0—10	5	1	–2	–2	4	–8	16
10—20	15	3	–1	–3	3	–3	3
20—30	25	4	0	0	0	0	0
30—40	35	2	+1	+2	2	+2	2
		$N = 10$		$\Sigma fd = -3$	$\Sigma fd^2 = 9$	$\Sigma fd^3 = -9$	$\Sigma fd^4 = 21$

$$\mu_1' = \frac{\Sigma fd}{N} \times i = \frac{-3}{10} \times 10 = -3$$

$$\mu_2' = \frac{\Sigma fd^2}{N} \times i^2 = \frac{9}{10} \times (10)^2 = \frac{9}{10} \times 100 = 90$$

$$\mu_3' = \frac{\Sigma fd^3}{N} \times i^3 = \frac{-9}{10} \times 1{,}000 = -900$$

$$\mu_4' = \frac{\Sigma fd^4}{N} \times i^4 = \frac{21}{10} \times 10{,}000 = 21{,}000$$

Moments about mean $\mu_1 = 0$

$$\mu_2 = \mu_2' - (\mu_1')^2 = 90 - (-3)^2 = 81$$

$$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2\,\mu_1'^3 = -900 - 3\,[(-3)\,(90)] + 3\,(-3)^3$$
$$= -900 + 810 - 54 = -144$$

$$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6\,(\mu_1')^2\,\mu_2' - 3\mu_1'^4$$
$$= 21{,}000 - 4\,(-3)\,(-900) + 6\,(-3)^2\,90 - 3\,(-3)^4$$
$$= 21{,}000 - 10{,}800 + 4{,}860 - 243 = 14{,}817.$$

Moments about Zero $v_1 = A + \mu_1'$ or the mean $= 25 - 3 = 22$

$$v_2 = \mu_2 + (v_1)^2 = 81 + (22)^2 = 565$$

$$v_3 = \mu_3 + 3v_1v_2 - 2v_1^3 = -144 + 3\,(22)\,(565) - 2\,(22)^3$$
$$= -144 + 37{,}290 - 21{,}296 = 15{,}850$$

$$v_4 = \mu_4 + 4v_1v_3 - 6v_1^2v_2 + 3v_1^4$$
$$= 14817 + 4\,(22)\,(15850) - 6\,(22)^2\,(565) + 3\,(22)^4$$
$$= 14817 + 1394800 - 1640760 + 702768 = 471625.$$

Illustration 11. The first four moments of a distribution about $x = 2$ are 1, 2.5, 5.5 and 16. Calculate the four moments about $\overline{X}$ and about zero.

(M. Com., Delhi Univ.; M. Com., M.D. Univ.; B. Com. (H), Delhi, 1999)

Solution. We are given $\mu_1' = 1$, $\mu_2' = 2.5$, $\mu_3' = 5.5$ and $\mu_4' = 16$. From these moments about arbitrary origin, we can find out moments about means with the help of the following relationship :

$$\mu_2 = \mu_2' - (\mu_1')^2$$
$$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2\,(\mu_1')^3$$
$$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6\,(\mu_1')^2\,\mu_2' - 3\,(\mu_1')^4$$

Substituting the value

$$\mu_2 = 2.5 - (1)^2 = 1.5$$
$$\mu_3 = 5.5 - 3\,(1)\,(2.5) + 2\,(1)^2 = 5.5 - 7.5 + 2 = 0$$
$$\mu_4 = 16 - 4\,(1)\,(5.5) + 6\,(2.5)\,(1)^2 - 3\,(1)^4 = 16 - 22 + 15 - 3 = 6.$$

Thus moments about mean are $\mu_1 = 0$, $\mu_2 = 1.5$, $\mu_2 = 0$, $\mu_4 = 6$.

Moment about zero :

Let moment about zero be denoted by v_1, v_2, v_3 etc.

The first moment about zero, *i.e.*, $v_1 = A + \mu_1'$ or mean

The second moment about zero *i.e.*, $v_2 = \mu_2 + v_1^2$

The third moment about zero *i.e.*, $v_3 = \mu_2 + 3v_1 v_2 - 2v_1^3$

The fourth moment about zero *i.e.*, $v_4 = \mu_1 + 4v_1 v_3 - 6v_1^2 v_2 + 3v_1^4$

In this question,

$v_1 = 2 + 1 = 3$; $\quad v_3 = 0 + (3 \times 3 \times 10.5) - 2\,(3)^3 = 94.5 - 54 = 40.5$

$v_2 = 1.5 + (3)^2 = 10.5$; $\quad v_4 = 6 + 4\,(3)\,(40.5) - 6\,(3)^2\,(10.5) + 3\,(3)^4$

$= 6 + 486 - 567 + 243 = 168.$

Measure of Skewness based on Moments

A measure of skewness is obtained by making use of the second and third moments about the mean. When the method of moments is applied β_1 is used as relative measure of skewness, β_1 is defined as :

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

In a symmetrical distribution β_1 shall be zero. The greater the value of β_1 the more skewed the distribution. 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_2 being the sum of the cubes of the deviations from the mean may be positive or negative but μ_3^2 is always positive. Also μ 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.*,

$$\gamma_1 = \sqrt{\beta_1} = \frac{\mu_3}{\mu_2^3} = \frac{\mu_3}{\sigma_3}$$

The sign of skewness would depend upon the value of μ_2. 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.

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 normal curve. In other words, measures of kurtosis tell us the extent to which a distribution is more peaked or flat-topped than the normal curve.* If a curve is more peaked than the normal curve, it is called '*leptokurtic*'. In such a case items are more closely bunched around the mode. On the other hand, if a curve is more flat-topped than the normal curve, it is called '*platykurtic*'. The normal curve itself is known as '*mesokurtic*'. The condition of peakedness or flat-toppedness itself is known as kurtosis of excess.** The concept of kurtosis is rarely used in elementary statistical analysis.

* "Kurtosis is the degree of peakedness of a distribbtion, usually taken relative to a normal distribution."–Spiegel : *Theory and Problems of Statistics*, p. 93.

** A.E. Waugh : *Elements of Statistical Methods*.