Pearson's Co-efficients

Karl Pearson defined the following

four coefficients, based upon the first four moments about mean :

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}, \quad \gamma_1 = +\sqrt{\beta_1} \quad \text{and} \quad \beta_2 = \frac{\mu_4}{\mu_2^2}, \quad \gamma_2 = \beta_2 - 3$$

Skewness

Literally, skewness means '*lack of symmetry*'. We study skewness to have an idea about the shape of the curve which we can draw with the help of the given data. A distribution is said to be skewed if

(*i*) Mean, median and mode fall at different points,

i.e., Mean $\neq$ Median $\neq$ Mode,

(*ii*) Quartiles are not equidistant from median, and

(*iii*) The curve drawn with the help of the given data is not symmetrical but stretched more to one side than to the other.

Measures of Skewness

Measures of Skewness. Various measures of skewness are

(1) $S_k = M - M_d$ (2) $S_k = M - M_0$.

where M is the mean, M_d, the median and M_0, the mode of the distribution.

(3) $S_k = (Q_3 - M_d) - (M_d - Q_1)$.

These are the absolute measures of skewness. As in dispersion, for comparing two series we do not calculate these absolute measures but we calculate the relative measures called the *co-efficients of skewness* which are pure numbers independent of units of measurement.

Co-efficients of Skewness

1. *Prof. Karl Pearson's Coefficient of Skewness.*

$$S_k = \frac{(M - M_0)}{\sigma}$$

II. *Prof. Bowley's Coefficient of Skewness.* Based on quartiles,

$$S_K = \frac{(Q_3 - M_d) - (M_d - Q_1)}{(Q_3 - M_d) + (M_d - Q_1)} = \frac{Q_3 + Q_1 - 2M_d}{Q_3 - Q_1}$$

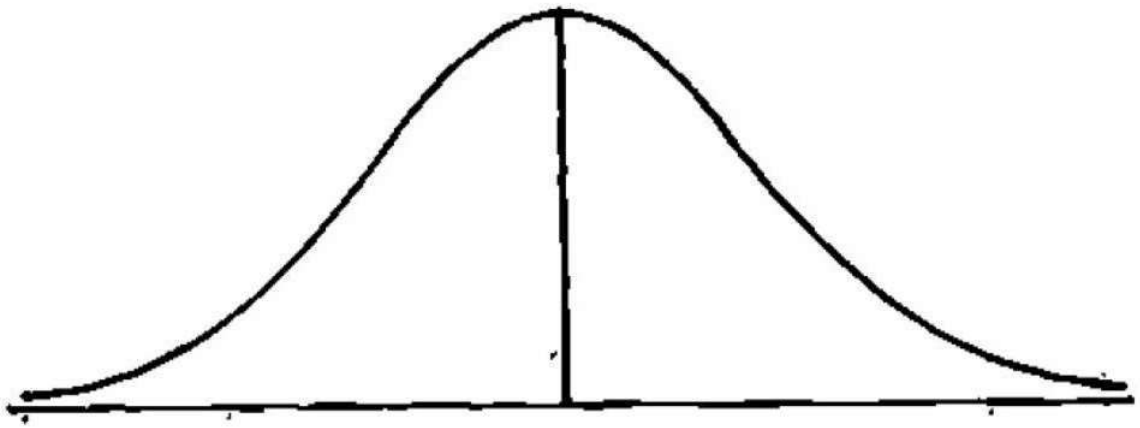
III. Based upon moments, co-efficient of skewness is

$$S_k = \frac{\sqrt{\beta_1}\ (\beta_2 + 3)}{2\ (5\beta_2 - 6\beta_1 - 9)}$$

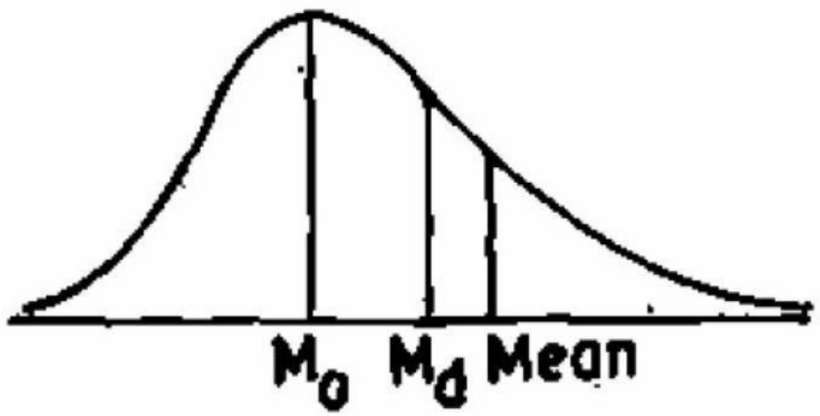
where symbols have their usual meaning. Thus $S_k = 0$ if either $\beta_1 = 0$ or $\beta_2 = -3$. But since $\beta_2 = \mu_4/\mu_2^2$, cannot be negative, $S_k = 0$ if and only if $\beta_1 = 0$. Thus for a symmetrical distribution $\beta_1 = 0$. In this respect β_1 is taken to be a measure of skewness.

The skewness is positive if the larger tail of the distribution lies towards the higher values of the variate (the right), i.e., if the curve drawn with the help of the given data is tretched more to the right than to the left and is negative in the contracy case.

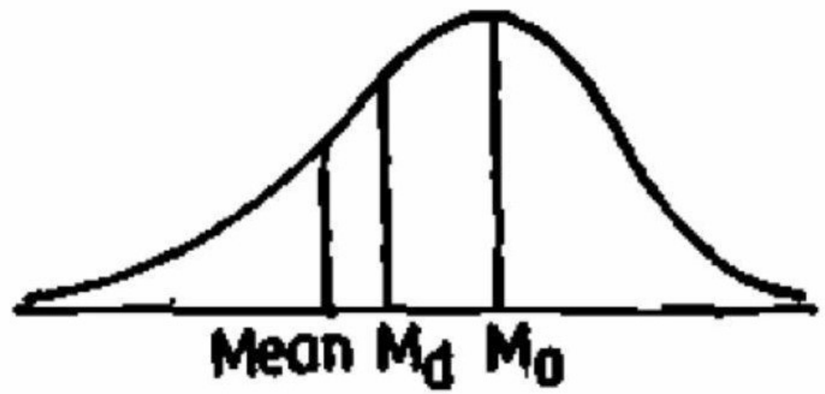
Graphical Representation of Skewness



$\bar{x}$ (*Mean*) = M_0 = M_d
(Symmetrical Distribution)



(Positively Skewed Distribution)



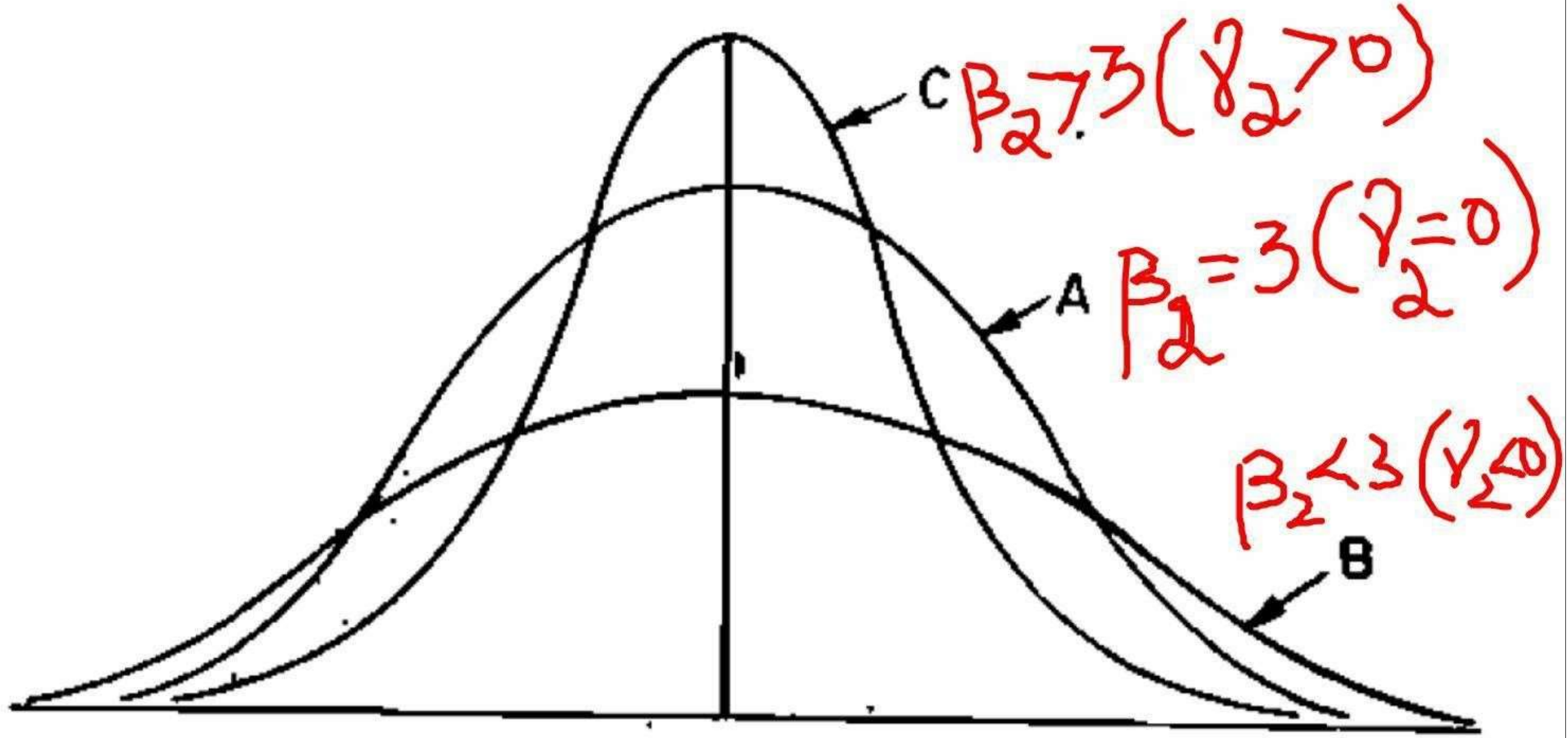
(Negatively Skewed Distribution)

Kurtosis

If we know the measures of central tendency, dispersion and skewness, we still cannot form a complete idea about the distribution as will be clear from the following figure in which all the three curves A, B and C are symmetrical about the mean 'm' and have the same range.

In addition to these measures we should know one more measure which Prof. Karl Pearson calls as the '*Convexity of curve*' or *Kurtosis.* Kurtosis enables us to have an idea about the flatness or peakedness of the curve. It is measured by the co-efficient β_2 or its derivation r_2 given by

$$\beta_2 = \mu_4/\mu_2^2, \quad \gamma_2 = \beta_2 - 3$$



Curve of the type '*A*' which is neither flat nor peaked is called the *normal curve* or *mesokurtic curve* and for such a curve $\beta_2 = 3$, i.e., $\gamma_2 = 0$. Curve of the type '*B*' which is flatter than the normal curve is known as *platykurtic* and for such a curve $\beta_2 < 3$, *i.e.*, $\gamma_2 < 0$. Curve of the type '*C*' which is more peaked than the normal curve is called *leptokurtic* and for such a curve $\beta_2 > 3$, i.e., $\gamma_2 > 0$.

Problems based on the Moments, Skewness and Kurtosis concepts:

Q1) Find the first four moments about $x=10$ for the series 4,7,10,13,16,19,22.

Q2) Calculate the first four moments about the mean for the series 4,7,10,13,16, 19,22.

Q3) The first four moments of a distribution about $x=4$ are 1, 4, 10 and 45. Comment upon the nature of the distribution.

Q4) In a certain distribution, the first four moments about x=5 are 2, 20, 40 and 50. Calculate and state whether the distribution is leptokurtic or platykurtic.

Q5) The first four central moments of a distribution are 0, 2.5, 0.7 and 18.75. Test the skewness and kurtosis of the distribution.