Axis 6: Form measurements (KEWNESS AND KURTOSIS)

Lack of symmetry is called skewness for a frequency distribution. If the distribution is not symmetric, the frequencies will not be uniformly distributed about the centre of the distribution. Here, we shall study various measures of skewness and kurtosis.
In this unit, the concepts of skewness are described in Section 4.2 whereas the various measures of skewness are given with examples in Section 4.3. In Section 4.4, the concepts and the measures of kurtosis are described.

Objectives

On studying this unit, you would be able to describe the concepts of skewness; explain the different measures of skewness; describe the concepts of kurtosis; explain the different measures of kurtosis; and explain how skewness and kurtosis describe the shape of a distribution.

I. CONCEPT OF SKEWNESS

Skewness means lack of symmetry. In mathematics, a figure is called symmetric if there exists a point in it through which if a perpendicular is drawn on the X-axis, it divides the figure into two congruent parts i.e. identical in all respect or one part can be superimposed on the other i.e mirror images of each other. In Statistics, a distribution is called symmetric if mean, median and mode coincide. Otherwise, the distribution becomes asymmetric. If the right tail is longer, we get a positively skewed distribution for which mean > median > mode while if the left tail is longer, we get a negatively skewed distribution for which mean < median < mode.
The example of the Symmetrical curve, Positive skewed curve and Negative skewed curve are given as follows:

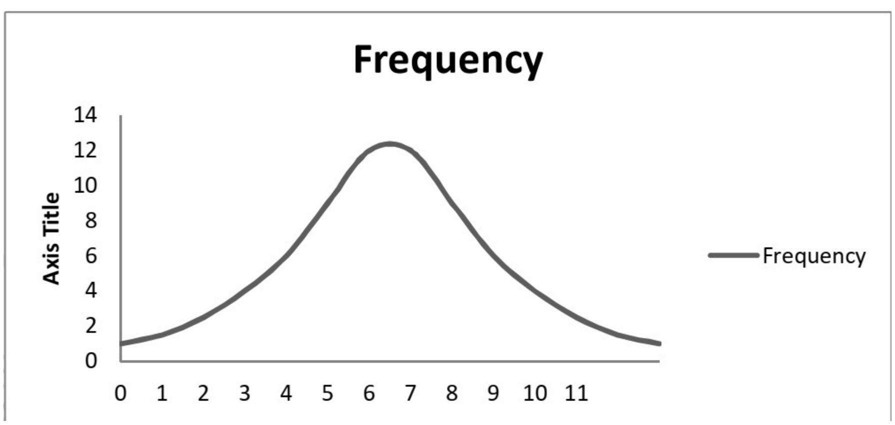

Fig. 1: Symmetrical Curve

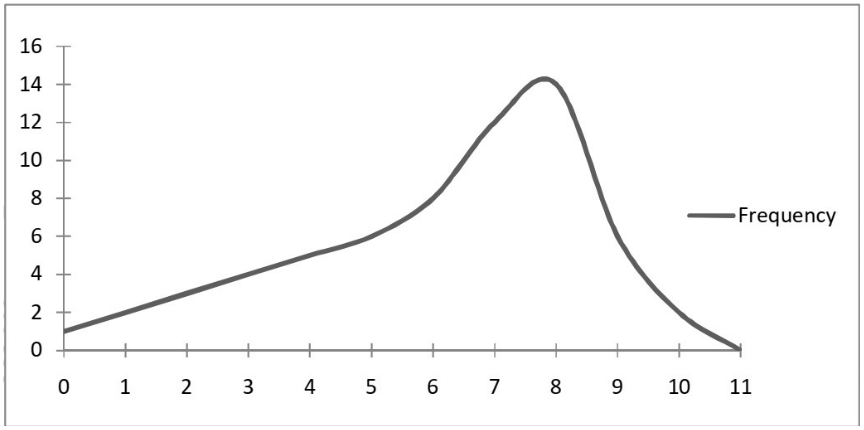

Fig. 2: Negative Skewed Curve

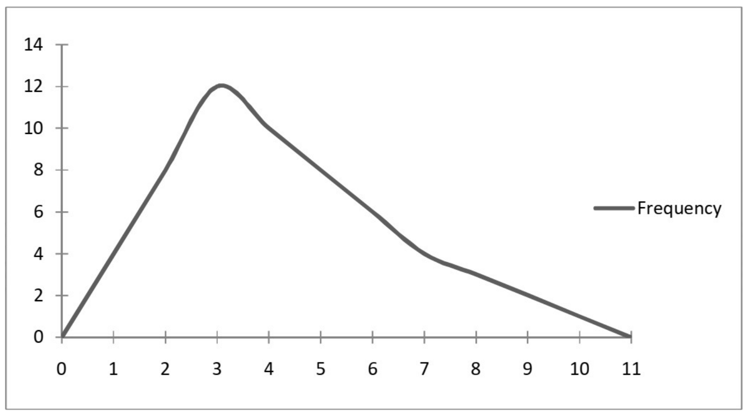

Fig. 3: Positive Skewed Curve

Difference between Variance and Skewness

The following two points of difference between variance and skewness should be carefully noted.

1. Variance tells us about the amount of variability while skewness gives the direction of variability.
2. In business and economic series, measures of variation have greater practical application than measures of skewness. However, in medical and life science field measures of skewness have greater practical applications than the variance.

II. VARIOUS MEASURES OF SKEWNESS

Measures of skewness help us to know to what degree and in which direction (positive or negative) the frequency distribution has a departure from symmetry. Although positive or negative skewness can be detected graphically depending on whether the right tail or the left tail is longer but, we don't get idea of the magnitude. Besides, borderline cases between symmetry and asymmetry may be difficult to detect graphically. Hence some statistical measures are required to find the magnitude of lack of symmetry. A good measure of skewness should possess three criteria:

1. It should be a unit free number so that the shapes of different distributions, so far as symmetry is concerned, can be compared even if the unit of the underlying variables are different;
2. If the distribution is symmetric, the value of the measure should be zero. Similarly, the measure should give positive or negative values according as the distribution has positive or negative skewness respectively; and

3. As we move from extreme negative skewness to extreme positive skewness, the value of the measure should vary accordingly. Measures of skewness can be both absolute as well as relative. Since in a symmetrical distribution mean, median and mode are identical more the mean moves away from the mode, the larger the asymmetry or skewness. An absolute measure of skewness can not be used for purposes of comparison because of the same amount of skewness has different meanings in distribution with small variation and in distribution with large variation.

1 . Absolute Measures of Skewness

Following are the absolute measures of skewness:

1. Skewness (Sk) = Mean – Median
2. Skewness (Sk) = Mean – Mode
3. Skewness (Sk) = (Q3 - Q2) - (Q2 - Q1)

For comparing to series, we do not calculate these absolute mearues we calculate the relative measures which are called coefficient of skewness. Coefficient of skewness are pure numbers independent of units of measurements.

2 . Relative Measures of Skewness

In order to make valid comparison between the skewness of two or more distributions we have to eliminate the distributing influence of variation. Such elimination can be done by dividing the absolute skewness by standard deviation. The following are the important methods of measuring relative skewness:

2.1. βand γCoefficient of Skewness

Karl Pearson defined the followingβand γcoefficients of skewness, based upon the second and third central moments:

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

It is used as measure of skewness. For a symmetrical distribution, $\beta_1$shall be zero. β_1 as a measure of skewness does not tell about the direction of skewness, i.e. positive or negative. Because μ_3 being the sum of cubes of the deviations from mean may be positive or negative but $\mu_2^3$is always positive. Also, μ_2 being the variance always positive.

Hence, β_1 would be always positive. This drawback is removed if we calculate Karl Pearson's Gamma coefficient $\gamma_1$which is the square root of β_1 i. e.

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

Then the sign of skewness would depend upon the value of μ_3 whether it is positive or negative. It is advisable to use γ_1 as measure of skewness.

2.2. Karl Pearson's Coefficient of Skewness

This method is most frequently used for measuring skewness. The formula for measuring coefficient of skewness is given by

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

The value of this coefficient would be zero in a symmetrical distribution. If mean is greater than mode, coefficient of skewness would be positive otherwise negative. The value of the Karl Pearson's coefficient of skewness usually lies between ± 1 for moderately skewed distubution. If mode is not well defined, we use the formula

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

By using the relationship

Mode = (3 Median – 2 Mean)

Here, $-3 \leq S_k \leq 3$.

In practice it is rarely obtained.

3. CONCEPT OF KURTOSIS

If we have the knowledge of the measures of central tendency, dispersion and skewness, even then we cannot get a complete idea of a distribution. In addition to these measures, we need to know another measure to get the complete idea about the shape of the distribution which can be studied with the help of Kurtosis. Prof. Karl Pearson has called it the "Convexity of a Curve". Kurtosis gives a measure of flatness of distribution.

The degree of kurtosis of a distribution is measured relative to that of a normal curve. The curves with greater peakedness than the normal curve are called **"Leptokurtic".** The curves which are more flat than the normal curve are called **"Platykurtic"**. The normal curve is called "**Mesokurtic**." The Fig.4 describes the three different curves mentioned above:

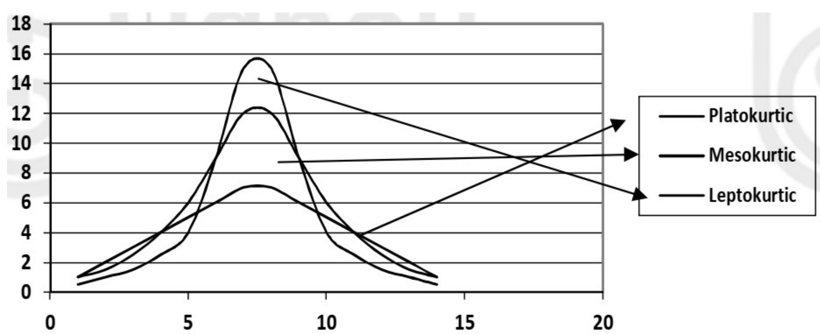

Fig.4: Platykurtic Curve, Mesokurtic Curve and Leptokurtic Curve.

4. Measures of Kurtosis

Karl Pearson's Measures of Kurtosis

For calculating the kurtosis, the second and fourth central moments of variable are used. For this, following formula given by Karl Pearson is used:

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

or $\gamma_2 = \beta_2 - 3$

where, μ_2 = Second order central moment of distribution

μ_4 = Fourth order central moment of distribution

Description:

1. If $\beta_2 = 3$ or $\gamma_2 = 0$, then curve is said to be mesokurtic;
2. If $\beta_2 < 3$ or $\gamma_2 < 0$, then curve is said to be platykurtic;
3. If $\beta_2 > 3$ or $\gamma_2 > 0$, then curve is said to be leptokurtic;