[Academic Script]

Moments, Skewness and Kurtosis

Subject:	Business Economics
Course:	B. A. (Hons.), 1st Semester, Undergraduate
Paper No. & Title:	Paper – 102 Statistics for Business Economics
Unit No. & Title:	Unit – 2 Univariate Analysis
Lecture No. & Title:	Lecture – 2 Moments, Skewness and Kurtosis

Academic Script

1. Moments

Moments are used to describe the peculiarities of a frequency distribution. Using moments, we can measure the central tendency of the data, their scatter i.e. dispersion, asymmetry and the peakedness of the curve.

	Moments	Central Moments (About Mean)	Raw Moments (About assumed mean A)
Ungrouped data	First	$\mu_1 = \frac{\sum(x-\bar{x})}{N}$	$\mu_1{}' = \frac{\sum(x-A)}{N}$
	Second	$\mu_2 = \frac{\sum(x-\bar{x})^2}{N}$	$\mu_2{}' = \frac{\sum(x-A)^2}{N}$
	Third	$\mu_3 = \frac{\sum(x-\bar{x})^3}{N}$	$\mu_3{}' = \frac{\sum(x-A)^3}{N}$
	Fourth	$\mu_4 = \frac{\sum(x-\bar{x})^4}{N}$	$\mu_4{}' = \frac{\sum(x-A)^4}{N}$
Grouped data	First	$\mu_1 = \frac{\sum f(x-\bar{x})}{N}$	$\mu_1{}' = \frac{\sum f(x-A)}{N}$
	Second	$\mu_2 = \frac{\sum f(x-\bar{x})^2}{N}$	$\mu_2{}' = \frac{\sum f(x-A)^2}{N}$
	Third	$\mu_3 = \frac{\sum f(x-\bar{x})^3}{N}$	$\mu_3{}' = \frac{\sum f(x-A)^3}{N}$
	Fourth	$\mu_4 = \frac{\sum f(x-\bar{x})^4}{N}$	$\mu_4{}' = \frac{\sum f(x-A)^4}{N}$

Note:

1. The first central moment is always zero i.e. $\mu_1 = 0$.

2. The second central moment indicates the variance.

Sheppard's correction

In case of grouped frequency, distribution while calculating moments we assume that the frequencies are concentrated at the centre of the class-intervals. However, the assumption is not true in practice and some error called the grouping error enters into the calculation of moments.

W. F. Sheppard's have given approximate corrections to estimates of moments .

The first few measured and corrected moments about the mean are then related as follows:

$$\mu_2\ (corrected) = \mu_2 - \frac{h^2}{12}$$

$$\mu_3\ (corrected) = \mu_3$$

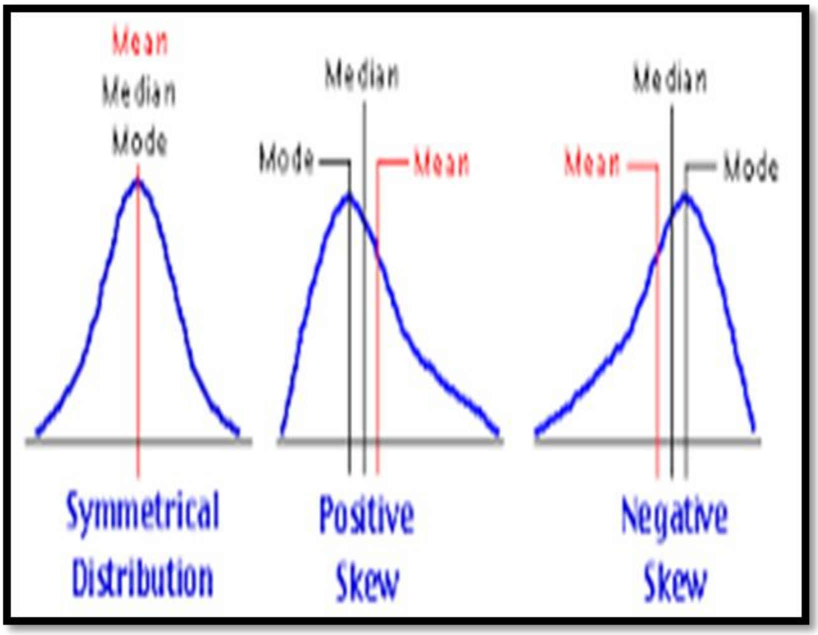
$$\mu_4\ (corrected) = \mu_4 - \frac{1}{2}h^2\mu_2 + \frac{7}{240}h^4$$

where h is the class interval

2. Skewness

A fundamental task in many of the statistical analysis is to characterize the location and variability of a data set. A further characterization of the data includes skewness and kurtosis.

Measure of Dispersion tells us about the variation of the data set. Skewness tells us about the direction of variation of the data set.



Definition: Skewness is a measure of symmetry, or more precisely, the lack of symmetry. A distribution, or data set, is symmetric if it looks the same on the left and right of the center point.

If in a distribution the value of the variable are distributed such that frequencies are equal, at equal distances, on either side of the central value, the distribution of the values is said to be symmetrical about the central value.

In the case of the symmetrical distribution, the two tails are of equal length; in the case of the asymmetrical distribution one tail is longer than the other. If the left tail is longer than the right tail, the distribution is said to be negatively skewed. If the right tail is longer than the left tail, the distribution is said to be positively skewed.

Measure of Skewness:

To find out the direction and the extent of asymmetry in the data, statistical measures of skewness are employed. These measures can be absolute or relative. The absolute measure of skewness tells us the extent of asymmetry and whether it is positive or negative. Relative measure is used for comparison.

Karl Pearson coefficient of Skewness	**Bowley's coefficient of Skewness**
$Sk. = Mean - Mode$ Coefficient of Skewness $= \frac{Mean - Mode}{Standard\ deviation}$ ***If mode is ill-defined*** $Sk. = 3(Mean - Median)$ Coefficient of Skewness $= \frac{3(Mean - Median)}{Standard\ deviation}$	$Sk. = Q_3 + Q_1 - 2M$ Coefficient of Skewness $= \frac{Q_3 + Q_1 - 2M}{Q_3 - Q_1}$

Skewness using moments:

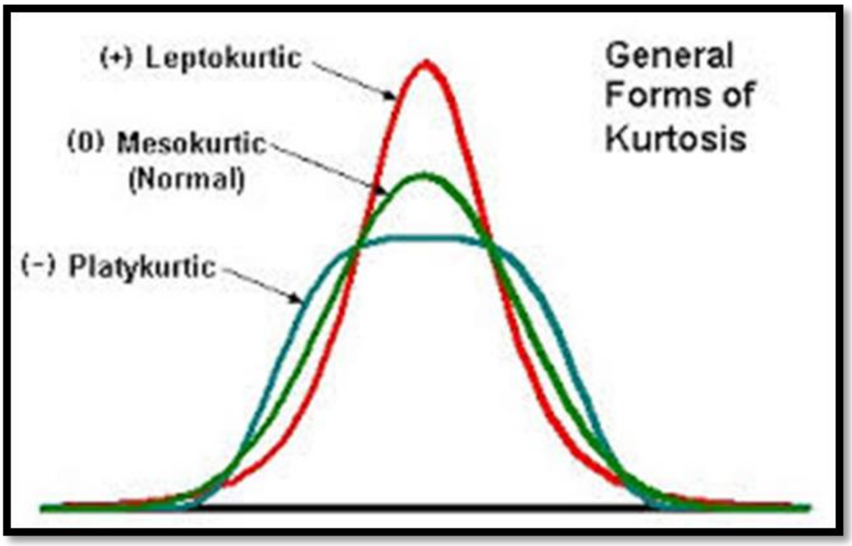
Karl Pearson as given measure of skewness using second and third central moments defined as

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

Coefficient of skewness = $\gamma_1 = \sqrt{\beta_1}$

3. Kurtosis

Kurtosis is a parameter that describes the shape of a random variable's probability distribution. Kurtosis characterizes the relative peakedness or flatness of a distribution compared to the normal distribution.



- A peaked curve is called "**Leptokurtic"** curve.
- An intermediate peaked curve which is neither flat-topped nor peaked is known as normal or "**Mesokurtic**" curve.
- A flat topped curve is called "**Platykurtic**" curve.

Karl Pearson as given measure of kurtosis using second and fourth central moments defined as

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

Coefficient of kurtosis = $\gamma_2 = \beta_2 - 3$

For a normal or mesokurtic curve, $\beta_2 = 3$ or $\gamma_2 = 0$.

For a leptokurtic curve $\beta_2 > 3$ or $\gamma_2 > 0$.

For a platykurtic curve $\beta_2 < 3$ or $\gamma_2 < 0$.

Example 2. The frequency distribution of heights of 100 college students is as follows:

Height (cm)	141 - 150	151 - 160	161 - 170	171 - 180	181 - 190
No. of students	5	16	56	19	4

Find skewness and kurtosis of the distribution using moments and interpret it.

Solution:

We shall obtain skewness and kurtosis using moments

Height	No. of students f	Mid value x	fx	$(x-\bar{x})$	$f(x-\bar{x})$	$f(x-\bar{x})^2$	$f(x-\bar{x})^3$	$f(x-\bar{x})^4$
141-150	5	145.5	727.5	-20.1	-100.5	2020.1	-40603.01	816120.40
151-160	16	155.5	2488	-10.1	-161.6	1632.2	-16484.82	166496.64
161-170	56	165.5	9268	-0.1	-5.6	0.56	-0.056	0.0056
171-180	19	175.5	3334.5	9.9	188.1	1862.2	18435.68	182513.24
181-190	4	185.5	742	19.9	79.6	1584	31522.40	627295.68
	100		**16560**		**0**	**7099**	**-7129.8**	**1792425.97**

$$\bar{x} = \frac{\sum fx}{\sum f}$$

$$= \frac{16560}{100}$$

$$\bar{x} = 165.60$$

First four central moments are:

$$\mu_1 = \frac{\sum f(x-\bar{x})}{N} = \frac{0}{100} = 0$$

$$\mu_2 = \frac{\sum f(x-\bar{x})^2}{N} = \frac{7099}{100} = 70.99$$

$$\mu_3 = \frac{\sum f(x-\bar{x})^3}{N} = \frac{-7129.8}{100} = -71.30$$

$$\mu_4 = \frac{\sum f(x-\bar{x})^4}{N} = \frac{1792425.97}{100} = 17924.26$$

Now, Skewness: $\beta_1 = \frac{{\mu_3}^2}{{\mu_2}^3}$

$= \frac{(-71.30)^2}{(70.99)^3}$

$= \frac{5083.69}{357759.79}$

= 0.014

Here μ_3 is negative so the given distribution is negatively skewed.

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

$= \frac{17924.26}{(70.99)^2}$

$= \frac{17924.26}{5039.58}$

=3.56

Here β_2 > 3, so it is leptokurtic curve.

4. Summary

- Moments is used to describe the peculiarities of a frequency distribution. Using moments, we can measure the central tendency of the data, their scatter i.e. dispersion, asymmetry and the peakedness of the curve.
- Skewness is a measure of symmetry, or more precisely, the lack of symmetry. A distribution, or data set, is symmetric if it looks the same to the left and right of the center point.
- A distribution can be positively skewed, negative skewed or symmetric.
- Kurtosis characterizes the relative peakedness or flatness of a distribution compared to the normal distribution.
- A distribution curve can be platykurtic, mesokurtic or leptokurtic.