Basic Statistics

Lesson 5: Moments: Skewness and Kurtosis

Skewness:

A frequency distribution is called symmetrical, when frequencies are symmetrically distributed about the mean of the distribution, i.e, values of the variables equidistant from mean have equal frequencies. Otherwise, the distribution is called asymmetrical or skew.

There are three types of frequency distributions:

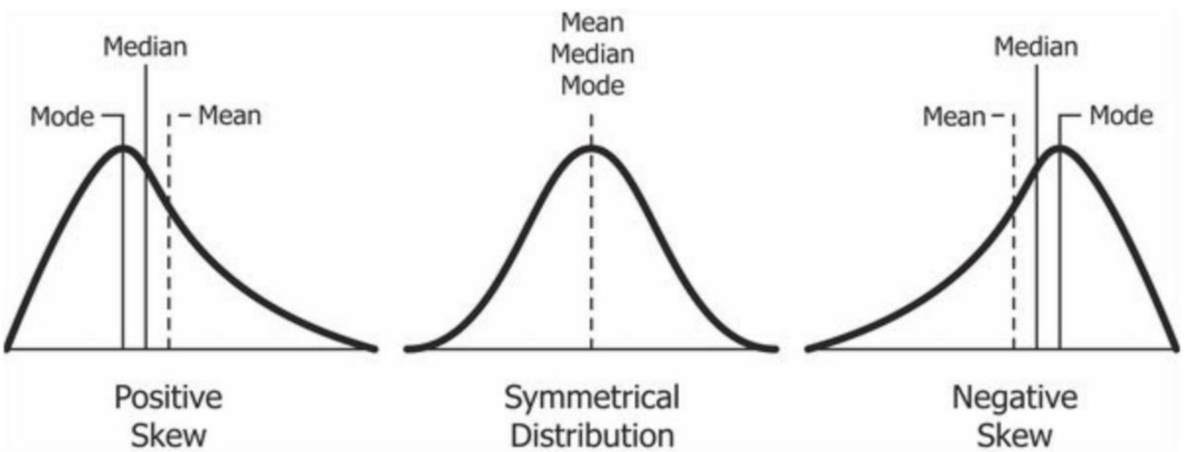


Fig. 1 — mean>median>mode

Fig. 2 — mean=median=mode

Fig. 3 — mean<median<mode

Skewness is measured by the following two ways:

(a) Karl Pearson's Coefficient of Skewness:

The measure is given by

$$S_k = \frac{\text{mean} - \text{mode}}{\text{standard deviation}}$$

If for some distributions, the mode is ill-defined, then

$$S_k = \frac{3(\text{mean} - \text{median})}{\text{standard deviation}}$$

Note: (i) $-1 \leq S_k \leq 1$

(ii) For comparing distributions, absolute values of S_k is to be considered

(iii) For Figs. (1), (2) and (3), $S_k > 0, = 0$ and < 0 respectively

(b) Moment Measure:

The skewness is defined as the third central moment of the distribution as

$$\gamma_1 = \frac{\mu_3}{\sigma^3}$$

where σ is the standard deviation of the distribution.

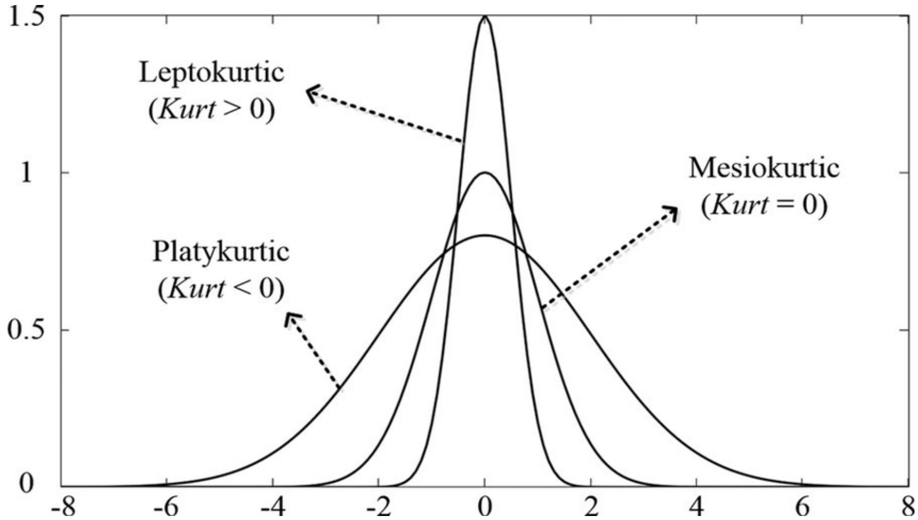
Kurtosis:

Kurtosis refers to the degree of "peakedness" or "tailedness" of a frequency distribution. The measure of this characteristic is given by the fourth central moment of the distribution as

$$\gamma_2 = \frac{\mu_4}{\sigma^4} - 3$$

According to the measurement of kurtosis, there are three types of frequency distributions:

- Leptokurtic ($\gamma_2 > 0$)
- Mesokurtic ($\gamma_2 = 0$)
- Platykurtic ($\gamma_2 < 0$)



Problems:

Ex.1. Find coefficient of skewness of the data given by:

3, 4, 8, 2, 4, 4, 3, 4, 7, 8, 9, 11

Solution: Let us make the table:

Obs. (x)	2	3	4	7	8	9	11	Total
Freq. (f)	1	2	4	1	2	1	1	12
xf	2	6	16	7	16	9	11	67
x^2	4	9	16	49	64	81	121	-
fx^2	4	18	64	49	128	81	121	465

The mean is given by

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i f_i}{N} = \frac{67}{12} = 5.58$$

The observation 4 has the highest frequency. Hence mode = 4.

Also the standard deviation is given by

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{n} f_i {x_i}^2 - \bar{x}^2} = \sqrt{\frac{465}{12} - 5.58^2} = 2.76$$

Then the coefficient of skewness is given by

$$S_k = \frac{\text{mean} - \text{mode}}{\text{standard deviation}} = \frac{5.58 - 4}{2.76} = 0.57$$

Ans.

Ex.2. Find the coefficient of skewness:

Marks	55-58	58-61	61-64	64-67	67-70
Freq.	12	17	23	18	11

Solution: Let us make the table for calculation of skewness:

Marks	55-58	58-61	61-64	64-67	67-70	Total
Freq.	12	17	23	18	11	81
x	56.5	59.5	62.5	65.5	68.5	-
xf	678	1011.5	1437.5	1179	753.5	5059.5
x^2	3192.25	3540.25	3906.25	4290.25	4692.25	-
fx^2	38307	60184.25	89843.75	77224.5	51614.75	317174.25

The mean is given by

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i f_i}{N} = \frac{5059.5}{81} = 62.46$$

The modal class is 61-64, hence the mode will be

$$\text{Mode} = l + \frac{h(f_1 - f_0)}{2f_1 - (f_0 + f_2)}$$

$$= 61 + \frac{3(23 - 17)}{46 - (17 + 18)} = 62.636$$

Also the standard deviation is given by

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{n} f_i {x_i}^2 - \bar{x}^2} = \sqrt{\frac{317174.25}{81} - 62.46^2}$$

$$= \sqrt{3915.73 - 3901.25} = \sqrt{14.48} = 3.8$$

Then the coefficient of skewness is given by

$$S_k = \frac{\text{mean} - \text{mode}}{\text{standard deviation}} = \frac{62.46 - 62.636}{3.8} = \frac{-0.176}{3.8} = -0.046$$

Ans.

Ex.3. A frequency distribution gives the following results: C.V = 5%, variance = 4, Karl Pearson's coefficient of skewness = 0.5. Find the mean and mode of the distribution.

Solution: We have C.V = 5% i.e, $5 = \frac{\sigma}{\bar{x}} \times 100 \;\rightarrow \bar{x} = \frac{2}{5} \times 100 \;\; [\text{variance} = 4]$

$= 40$

Also,

$$S_k = \frac{\text{mean} - \text{mode}}{\text{standard deviation}}$$

$$\text{or}, 0.5 = \frac{40 - \text{mode}}{2}$$

$$\text{or}, \text{mode} = 39$$

Ans.