# Skewness

Computer Oriented Numerical and Statistical Methods

Minal Shah

## Outline

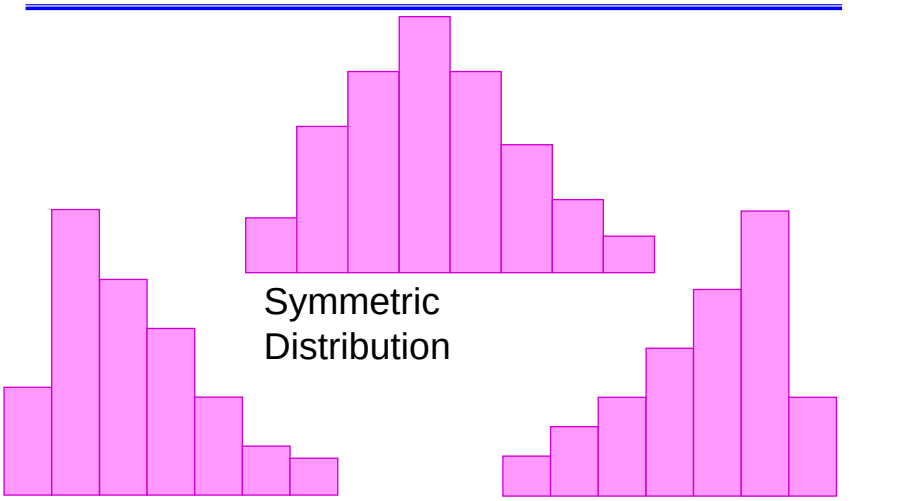
- Introduction
- Types of Skewness
- Measure of skewness.
- Karl Pearson's Measure
- Bowley's Measure

## Introduction

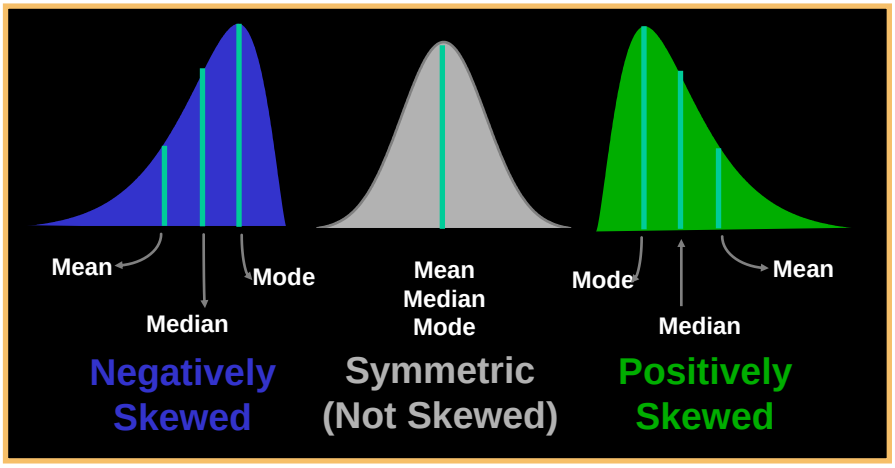
- Measure of central tendency gives us an estimate of the representative value of a series, the measure of dispersion gives an indication of the extent to which the items cluster around or scatter away from the central value and the skewness is a measure that refers to the extent of symmetry or asymmetry in a distribution.
- It describes the shape of a distribution,



## Skewness In The Form Of Histogram



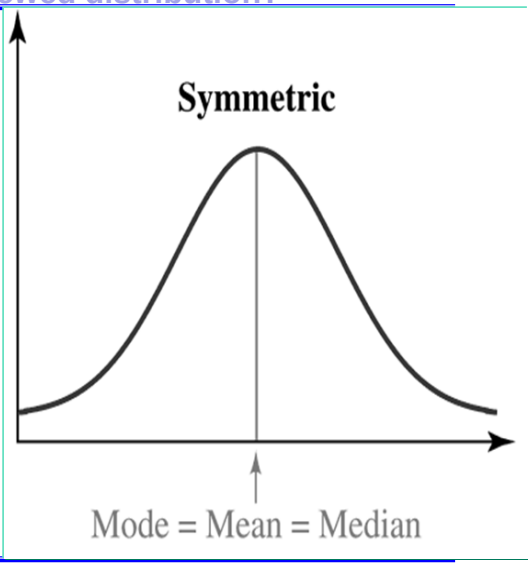
## Skewness





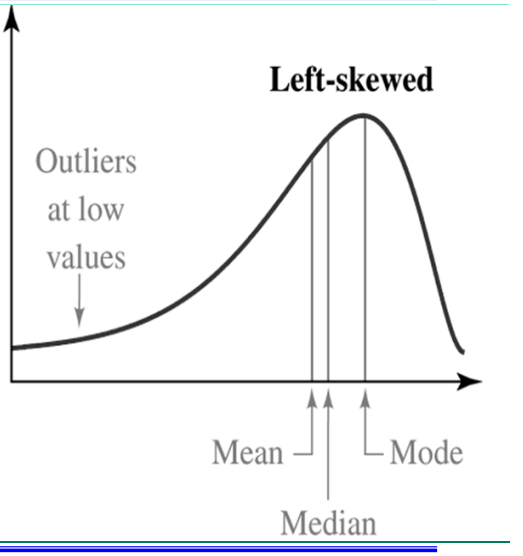

## What is the relationship between mean, median and mode of a skewed distribution?

- Find the mean median and mode of: 1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4, 4, 5, 5, 5, 6, 6, 7
- Mean is 4.
- Median is 4.
- Mode is 4.



## What is the relationship between mean, median and mode of a left-skewed distribution?

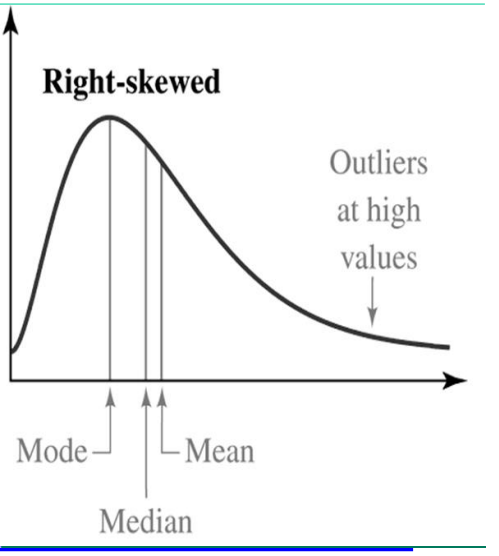
- Find the mean, median and mode of:
  0, 5, 10, 20, 40, 45, 45, 50, 50, 50, 60, 60, 60, 60, 60, 60, 70, 70, 70, 70, 70, 70, 70, 70
- The mean is 51.5.
- The median is 60.
- The mode is 70.






## What is the relationship between mean, median and mode of a right-skewed distribution?

- Find the mean, median, and mode of:
  20, 20, 20, 20, 20, 20, 20, 20, 30, 30, 30, 30, 30, 30, 45, 45, 45, 50, 50, 60, 70, 90
- The mean is 36.1.
- The median is 30.
- The mode is 20.



## Various Distribution And The Position Of Average

| Size | Frequency (c) | Frequency (b) | Frequency (a) |
|---|---|---|---|
| 0 – 5 | 10 | 10 | 10 |
| 5 – 10 | 90 | 30 | 20 |
| 10 – 15 | 50 | 50 | 30 |
| 15 – 20 | 40 | 70 | 40 |
| 20 – 25 | 30 | 50 | 50 |
| 25 – 30 | 20 | 30 | 90 |
| 30 – 35 | 10 | 10 | 10 |
| skewness | positive | symmetry | Negative |
| average | $\bar{x} > M_d > M_0$ | $\bar{x} = M_d = M_0$ | $\bar{x} < M_d < M_0$ |
| Quartiles | $Q_3 - M_d > M_d - Q_1$ | $Q_3 - M_d = M_d - Q_1$ | $Q_3 - M_d < M_d - Q_1$ |
| curve | Skewed to the right | Normal | Skewed to the left |



## Objectives of Skewness

- It helps in finding out the nature and the degree of concentration whether it is in higher or the lower values.
- The empirical relations of mean and median and mode are based on a moderately skewed distribution. The measure of skewness will reveal to what extent such empirical relationship holds good.
- It helps in knowing if the distribution is normal. Many statistical measures, such as the error of the mean are based on the assumption of a normal distribution.

## Measures of Skewness

- To find out the direction and the extent of asymmetry in a series statistical measures of skewness are employed.
- This measure can be absolute or relative.
- Absolute measure of skewness tell us the extent of asymmetry and whether it is positive or negative.
- The absolute skewness is based on the difference between mean and mode. Symbolically,

  absolute $S_k$ = Mean – Mode.
- If the value of mean is greater than the mode, skewness will be positive.



## Measures of Skewness

- If the value of mean is less than the mode, skewness will be negative.
- Absolute measure of skewness is not adequate because it cannot be used for comparison of skewness in two distributions if they are in different units, since difference between mean and mode will be in terms of units of distribution.
- For comparison purpose we use the relative measure of skewness known as coefficient of skewness.

## Measures of Skewness

- There are four types of relative measure of skewness :
  1. The Karl Pearson's Coefficient of Skewness.
  2. The Bowley's Coefficient of Skewness.
  3. The Kelly's Coefficient of Skewness.
  4. Measure of Skewness based on Moments and Kurtosis.



## Karl Pearson's Coefficient of Skewness

- Karl Pearson's Coefficient of Skewness or Pearsonian Coefficient of skewness is given by the formula:

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

- If in a particular frequency distribution, it is difficult to determine precisely the mode, or the mode is ill-defined, the coefficient of skewness can be determined by the following formula:

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

## Karl Pearson's Coefficient of Skewness

- Theoretically, skewness lies between the limits $\pm$ 3, but these limits are rarely attained in practice.



## Bowley's Coefficient of Skewness

- Bowley's coefficient of skewness also known as Quartile coefficient of skewness and is especially useful :
  - When the mode is ill-defined and extreme observations are present in the data.
  - When the distribution has open-end classes or unequal class-interval.
- The quartile measure depends upon the fact that normally $Q_3$ and $Q_1$ are equidistance from the median, i.e. for symmetrical distribution $Q_3 - M_d = M_d - Q_1$.
- If a distribution is asymmetrical, then one quartile will be farther from the median than the other.

## Bowley's Coefficient of Skewness

- In such a case skewness can be measured by the following formula given by Bowley :

  Skewness = $(Q_3 - M_d) - (M_d - Q_1)$

  Skewness = $Q_3 + Q_1 - 2M_d$
- If the first part is more than the second part, the skewness is positive and in the reverse situation it is negative.



## Bowley's Coefficient of Skewness

- To make the measure a readily comparable, the coefficient of skewness is obtained by dividing it by quartile range viz $Q_3 - Q_1$

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

$$S_k = \frac{Q_3 + Q_1 - 2M_d}{Q_3 - Q_1}$$

## Bowley's Coefficient of Skewness

- The range of variation under this method is $\pm 1$.
- The main drawback of this measure is that it is based on the central 50% of the data and ignores the remaining 50% of the data towards the extremes.