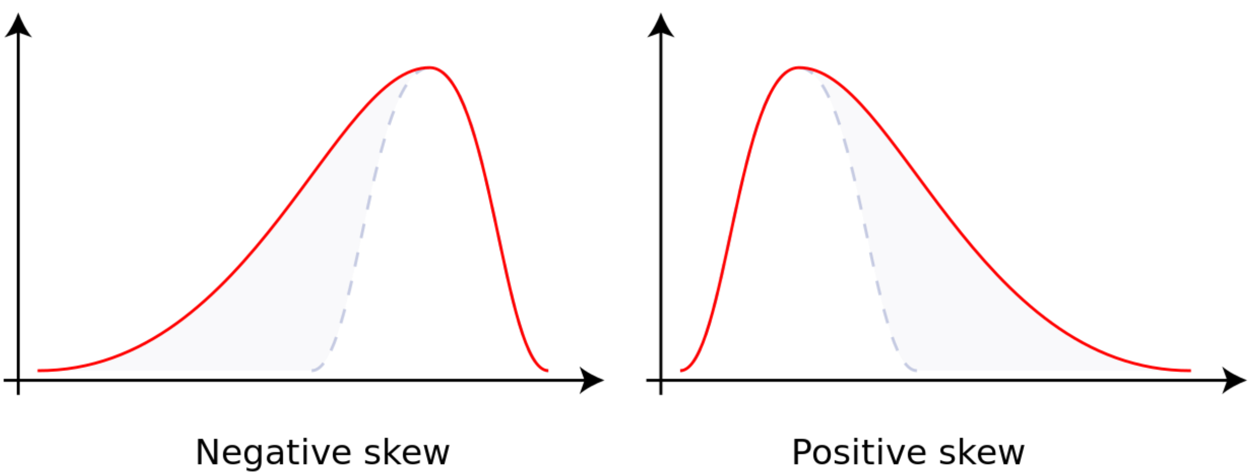
L6 Skewness

It is a measure of the asymmetry of the probability distribution of a real-valued random variable about its mean. The skewness value can be positive, zero, negative, or undefined.

1- **Negative skew:** The left tail is longer; the mass of the distribution is concentrated on the right of the figure. The distribution is said to be left-skewed, left-tailed, or skewed to the left, despite the fact that the curve itself appears to be skewed or leaning to the right; left instead refers to the left tail being drawn out and, often, the mean being skewed to the left of a typical center of the data. A left-skewed distribution usually appears as a right-leaning curve. Mean < Median < Mode

2- **Positive skew:** The right tail is longer; the mass of the distribution is concentrated on the left of the figure. The distribution is said to be right-skewed, right-tailed, or skewed to the right, despite the fact that the curve itself appears to be skewed or leaning to the left; right instead refers to the right tail being drawn out and, often, the mean being skewed to the right of a typical center of the data. A right-skewed distribution usually appears as a left-leaning curve. Mode < Median < Mean

- **zero value** in skewness means that the tails on both sides of the mean balance out overall; this is the case for a symmetric distribution. Mean= median=mode are equal then the distribution is a **normal distribution** and the coefficient of skewness will be 0.
- asymmetric distribution where one tail is long and thin, and the other is short but fat. Thus, the judgement on the symmetry of a given distribution by using only its skewness is risky; the distribution shape must be taken into account.



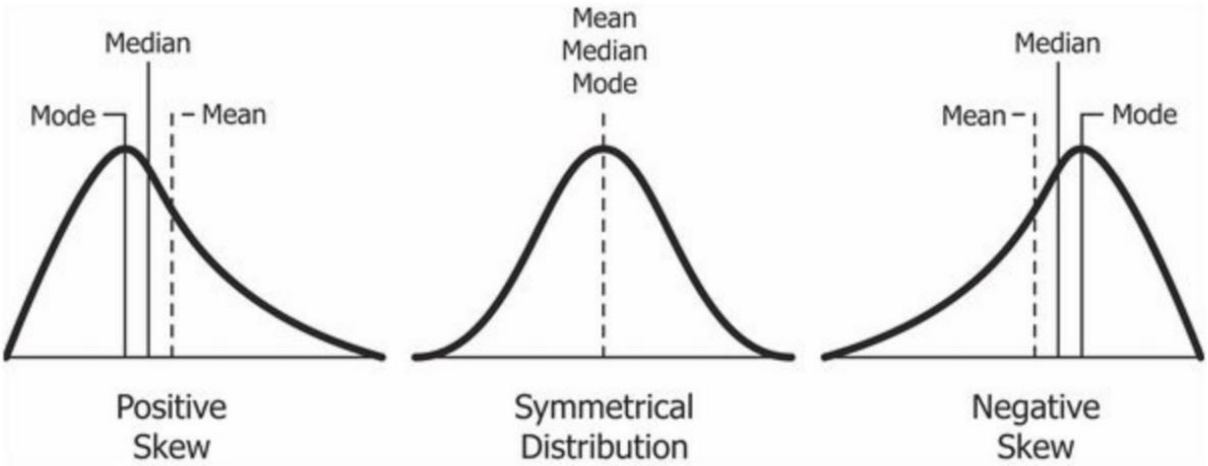
Nonparametric skew: defined as where is the mean, is the median, and is the standard deviation.

The skewness is defined in terms of this relationship:

- positive/right nonparametric skew means the mean is greater than (to the right of) the median,
- while negative/left nonparametric skew means the mean is less than (to the left of) the median.

If the distribution is symmetric, then the mean is equal to the median, and the distribution has zero skewness.

If the distribution is both symmetric and unimodal, then the mean = median = mode. This is the case of a coin toss or the series 1,2,3,4,...



Coefficient of Skewness Formula

Coefficient of Skewness

Using Mode: $\frac{\bar{x} - \text{Mode}}{s}$

Using Median: $\frac{3(\bar{x} - \text{Median})}{s}$

How to Calculate Coefficient of Skewness?

Depending upon the data available either of the two formulas can be used to calculate the coefficient of skewness.

EX: the mean of a data set is 60.5, the mode is 75, the median is 70 and the standard deviation is 10. The steps to calculate the coefficient of skewness are as follows:

Using Mode

- Step 1: Subtract the mode from the mean. mean -mode = 60.5 - 75 = -14.5
- Step 2: Divide this value by the standard deviation to get the coefficient of skewness. Thus, sk_1 = -14.5 / 10 = -1.45.

Using Median : the most popular called (Karl Pearson Coefficient of Skewness formula)

- Step 1: Subtract the median from the mean. mean- mode = 60.5 - 70 = -9.5

- Step 2: Multiply this value by 3. This gives -28.5.
- Step 3: Divide the value from step 2 by the standard deviation to obtain the coefficient of skewness. Thus, sk_2 = -28.5 / 10 = -2.85