Moments

In statistics, moments are a fundamental set of measures used to describe the shape and characteristics of a data distribution. They are mathematical quantities that summarize various aspects of a probability distribution, including its central location, spread, asymmetry, and tail behavior. In simple terms, a moment indicates how concentrated or dispersed the values are in relation to this central value. The most commonly used moments are the first four: the first moment is the mean, which represents the central tendency; the second moment is the variance, which measures dispersion; the third moment is the skewness, which indicates asymmetry; and the fourth moment is the kurtosis, which describes the peakedness and tail behavior of the distribution. Together, these moments provide a comprehensive summary of a dataset's statistical properties.

There are two main types of moments: raw moments and central moments. Raw moments (or moments about the origin) are calculated with respect to zero and are defined as the average of the values raised to a certain power. Central moments, on the other hand, are calculated with respect to the mean and measure how data values deviate from the average.

Raw moments	Central moments
$\mu_1' = \frac{\sum (x_i - A)}{n}$ $\mu_2' = \frac{\sum (x_i - A)^2}{n}$ $\mu_3' = \frac{\sum (x_i - A)^3}{n}$ $\mu_4' = \frac{\sum (x_i - A)^4}{n}$	$\mu_1 = \frac{\sum (x_i - \bar{x})}{n}$ $\mu_2 = \frac{\sum (x_i - \bar{x})^2}{n}$ $\mu_3 = \frac{\sum (x_i - \bar{x})^3}{n}$ $\mu_4 = \frac{\sum (x_i - \bar{x})^4}{n}$

Relationship between raw moments and central moments

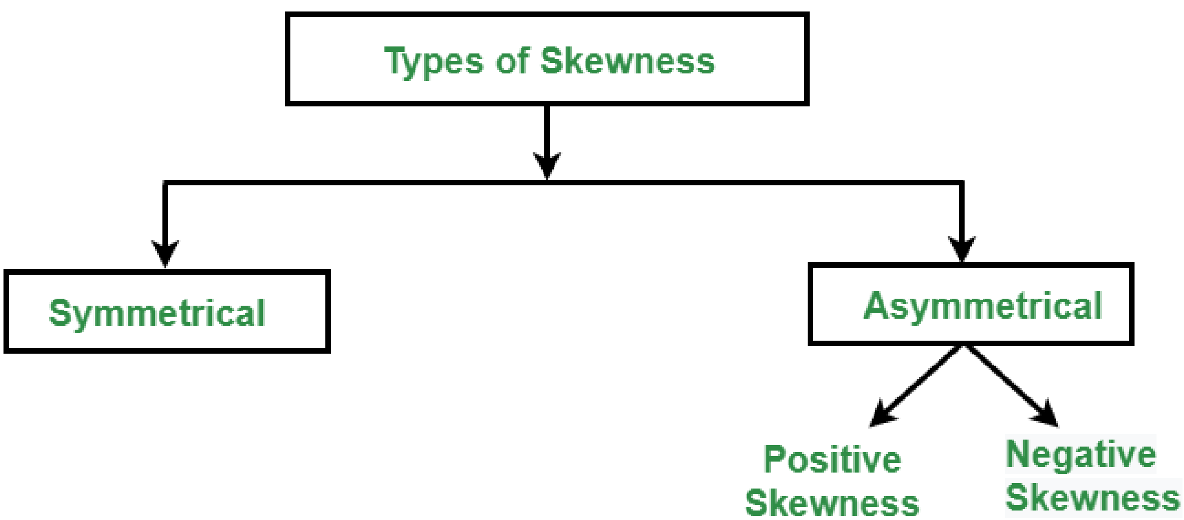
$$\mu_0 = \mu_0' = 1$$
$$\mu_1 = 0$$
$$\mu_2 = \mu_2' - \mu_1'^2$$
$$\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2\mu_1'^3$$
$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\mu_1'^2 - 3\mu_1'^4$$

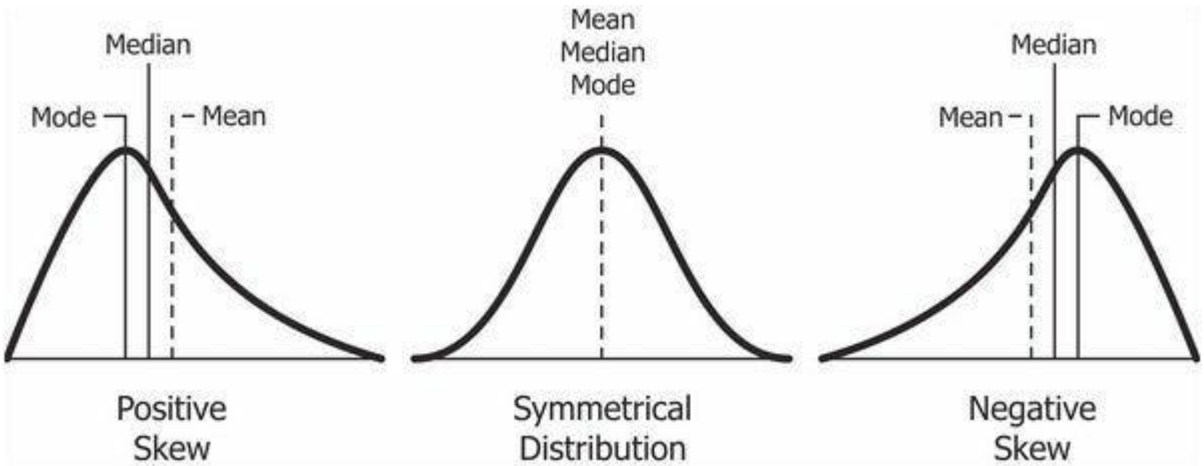
Skewness

Skewness is a statistical measure that describes the degree of asymmetry in a probability distribution around its mean. It indicates whether the data values tend to cluster more on one side of the mean than the other, showing if the distribution leans to the left or right.

Types of Skewness

1. **Positive Skewness (Right-skewed):** The distribution has a longer or fatter tail on the right side. Most data points lie to the left of the mean, but there are some extreme high values stretching the tail to the right.
2. **Negative Skewness (Left-skewed):** The distribution has a longer or fatter tail on the left side. Most data points lie to the right of the mean, but some extreme low values extend the tail to the left.
3. **Zero Skewness (Symmetric):** The distribution is perfectly symmetrical around the mean, with tails of equal length on both sides.





Decision Rule for Skewness

- If skewness is close to zero (commonly between -0.5 and +0.5), the distribution is considered approximately symmetric.
- Skewness values between 0.5 and 1 indicate moderate positive skewness; values greater than 1 imply high positive skewness.
- Skewness values between -0.5 and -1 indicate moderate negative skewness; values less than -1 imply high negative skewness.

Coefficient of Skewness

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

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

$$\text{skewness} = \frac{m_3}{m_2^{3/2}}$$

where

$$m_2 \text{ (the second sample moment about mean)} = \frac{\sum(x-\overline{x})^2}{n}$$

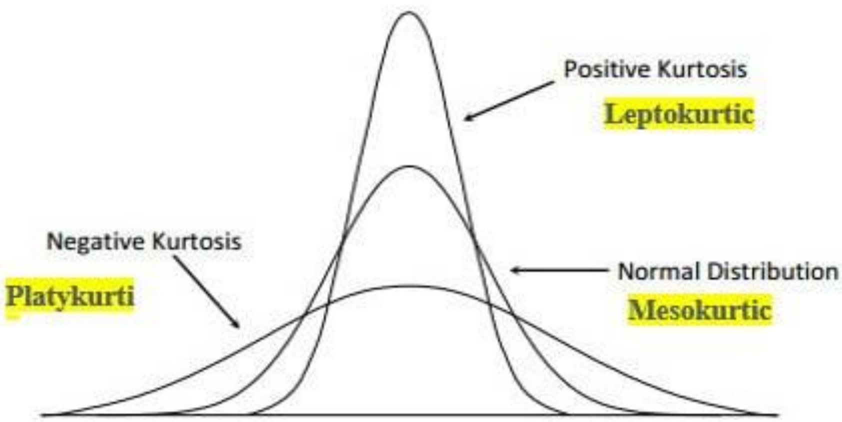
$$m_3 \text{ (the third sample moment about mean)} = \frac{\sum(x-\overline{x})^3}{n}$$

Kurtosis

Kurtosis is a statistical measure that describes the tailedness or peakedness of a probability distribution. It indicates how much of the variance in the data is due to extreme values (outliers) compared to a normal distribution. In other words, kurtosis measures whether the data have heavier or lighter tails and how sharp or flat the peak is.

Types of Kurtosis

1. **Mesokurtic:**
   A distribution with kurtosis equal to 3 (excess kurtosis = 0) is called mesokurtic. The normal distribution is the classic example. It has a moderate peak and tails that are neither too heavy nor too light.
2. **Leptokurtic:**
   Distributions with kurtosis greater than 3 (excess kurtosis > 0) are called leptokurtic. They have sharper peaks and heavier tails, meaning more frequent extreme values (outliers) than a normal distribution.
3. **Platykurtic:**
   Distributions with kurtosis less than 3 (excess kurtosis < 0) are called platykurtic. They have flatter peaks and lighter tails, indicating fewer and less extreme outliers than a normal distribution.



$$kurtosis = \frac{m_4}{m_2^2} = \frac{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4\right)}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^2}$$

$$skewness = \frac{(x - \mu)^3}{\sigma^3}$$

$$kurtosis = \frac{(x - \mu)^4}{\sigma^4}$$