Skewness and Kurtosis

Kaustubh Kulkarni

Skewness

- In everyday language, the terms “skewed” and “askew” are used to refer to something that is out of line or distorted on one side.
- When referring to the shape of frequency or probability distributions, “skewness” refers to asymmetry of the distribution.
- A distribution with an asymmetric tail extending out to the right is referred to as “**positively skewed**” or “**skewed to the right**”.
- A distribution with an asymmetric tail extending out to the left is referred to as “**negatively skewed**” or “**skewed to the left**.”
- Skewness can range from minus infinity to positive infinity.

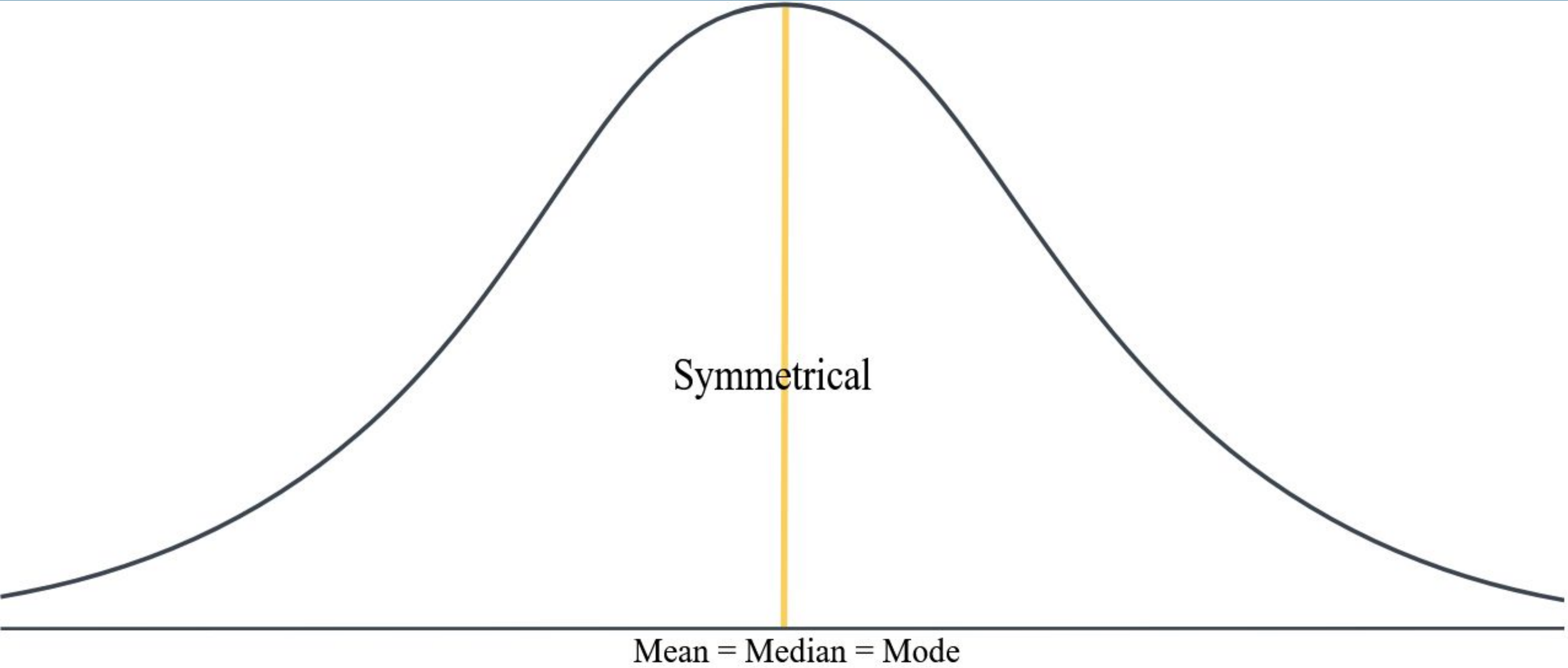
Skewness Formula

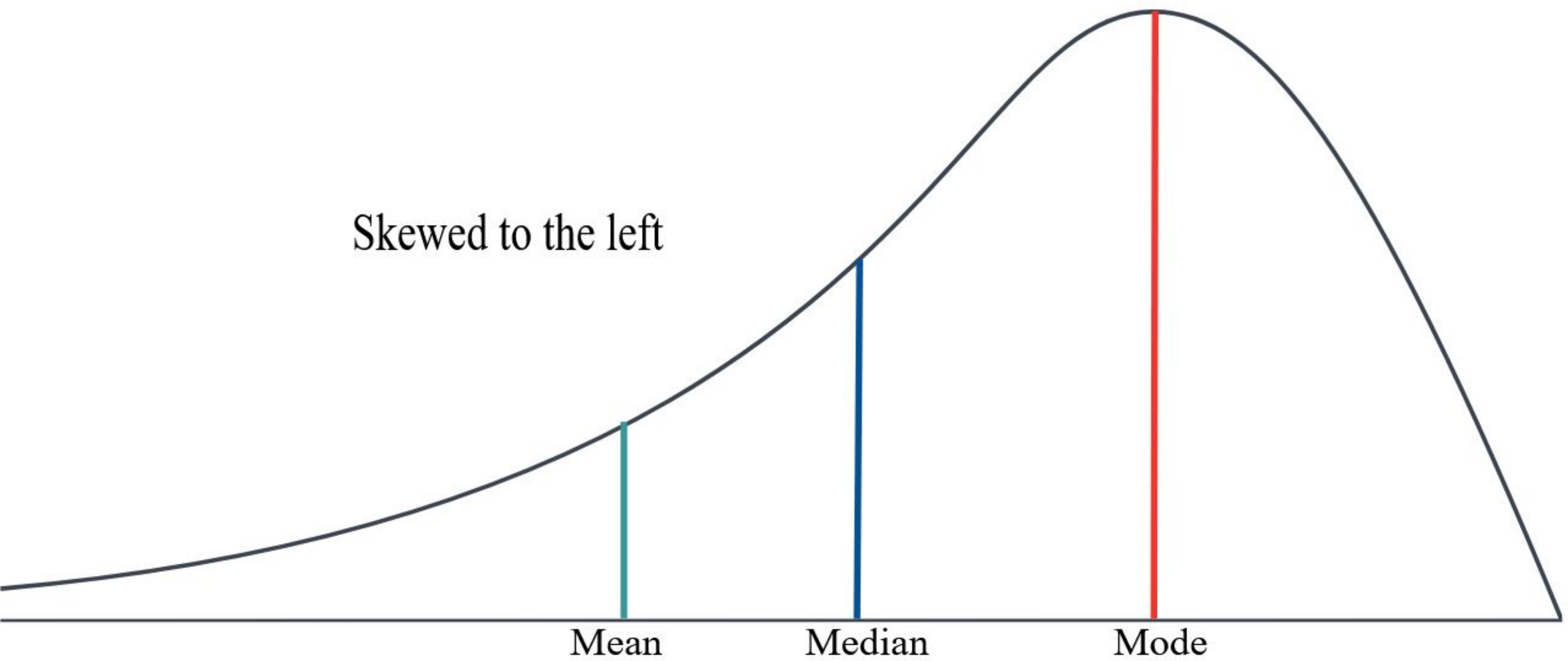
- Karl Pearson (1895) first suggested measuring skewness by standardizing the difference between the mean and the mode:

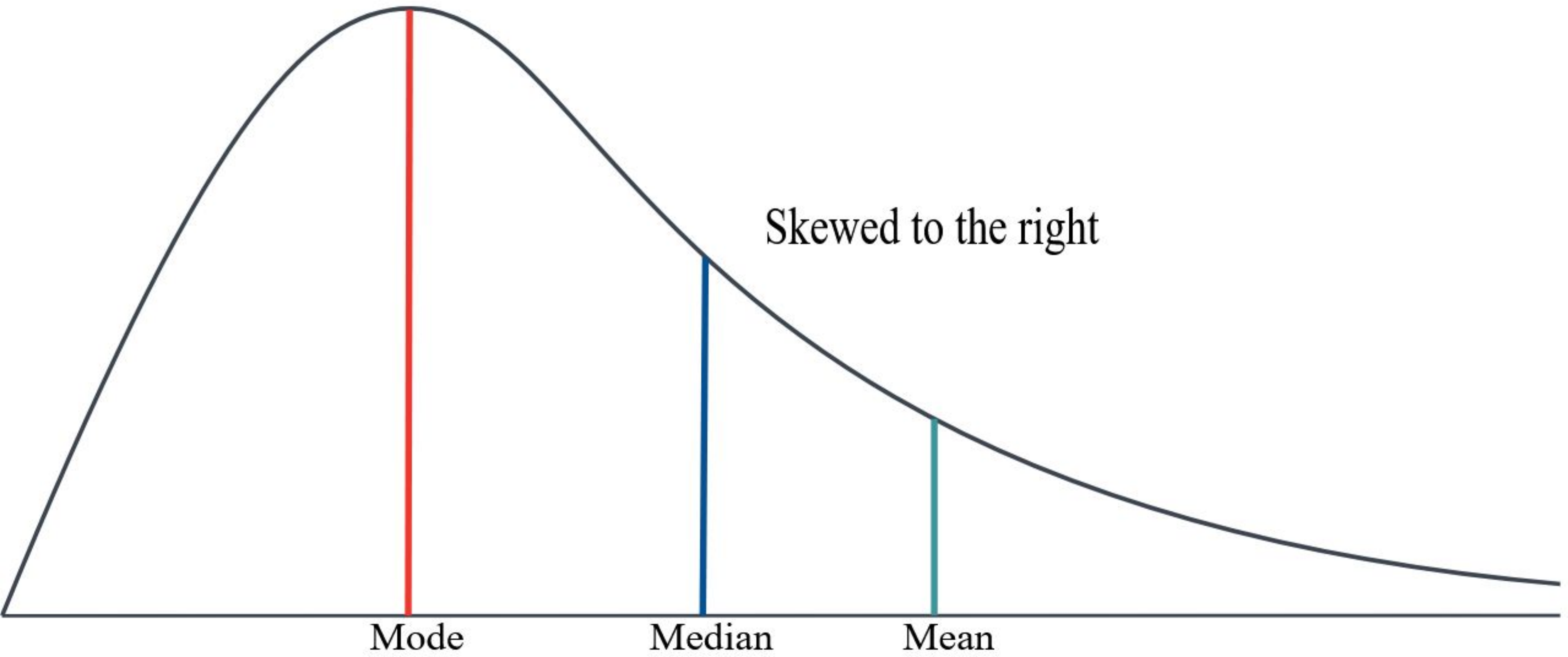
$$sk = \frac{\mu - \text{mode}}{\sigma}$$

- Population modes are not well estimated from sample modes
- Difference between the mean and the mode as being three times the difference between the mean and the median (Stuart & Ord, 1994):

$$sk_{est} = \frac{3(M - \text{median})}{s}$$


Symmetrical
Mean = Median = Mode


Skewed to the left
Mean
Median
Mode


Skewed to the right
Mode
Median
Mean

Kurtosis

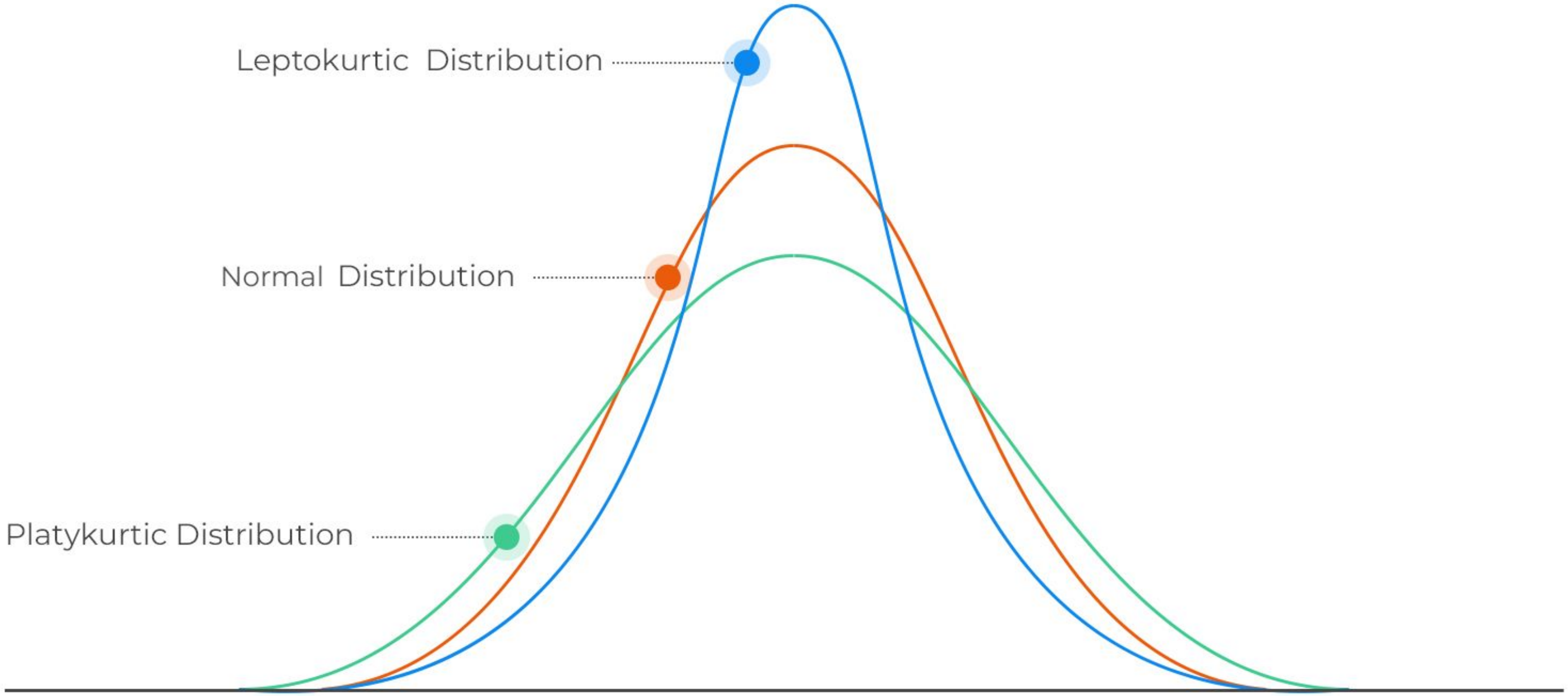
Kurtosis refers to measuring the degree to which a given distribution is more or less ‘peaked’ relative to the normal distribution.

The concept of kurtosis is very useful in decision-making.

In this regard, we have 3 categories of distributions:

1. Leptokurtic
2. Mesokurtic
3. Platykurtic

Kurtosis


Leptokurtic Distribution
Normal Distribution
Platykurtic Distribution

Leptokurtic

- A leptokurtic distribution is more peaked than the normal distribution.
- The higher peak results from the clustering of data points along the X-axis.
- “Lepto” means thin.
- The tails are also fatter than those of a normal distribution.
- The coefficient of kurtosis is usually found to be more than 3.

Interpretation

- When analyzing historical returns, a leptokurtic distribution means that changes are less frequent since historical values are clustered around the mean.
- However, there are also large fluctuations represented by the fat tails.

Platykurtic

- A platykurtic distribution has extremely dispersed points along the X-axis, resulting in a lower peak when compared to the normal distribution.
- “Platy” means broad.
- Hence, the prefix fits the distribution’s shape, which is wide and flat.
- The points are less clustered around the mean compared to the leptokurtic distribution.

Platykurtic

- Returns that follow this type of distribution have fewer major fluctuations compared to leptokurtic returns.
- However, you should note that fluctuations represent the riskiness of an asset.
- More fluctuations represent more risk and vice versa.
- Therefore, platykurtic returns are less risky than leptokurtic returns.

Mesokurtic

- Meso = middle; intermediate.
- This means the distribution is a normal distribution.

Sample Skewness

$$S_k = \frac{1}{n} \frac{\sum_{i=1}^{n} (X_i - \bar{X})^3}{S^3}$$

Where: $\bar{X}$ is the sample mean;

S is the sample standard deviation; and

n is the number of observations.

Sample Kurtosis

Sample kurtosis is always measured relative to the kurtosis of a normal distribution, which is 3. Therefore, we are always interested in the "excess" kurtosis, i.e.,

Excess kurtosis = sample kurtosis – 3, where:

$$S_{kr} = \frac{1}{n}\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^4}{s^4}$$

Interpretation:

A positive excess kurtosis indicates a leptokurtic distribution.

A zero value indicates a mesokurtic distribution.

Lastly, a negative excess kurtosis represents a platykurtic distribution.

Suppose we have the following observations:

{12   13   54   56   25}

Determine the skewness and kurtosis of the data.

Raw Moment of a PDF

Definitions of raw moments

For a continuous probability distribution for density function f(x), the nth raw moment (also called the moment about zero) is defined as

$$\mu'_n = \int_{-\infty}^{\infty} x^n f(x)\, dx$$

Observation

The mean can be defined as the first raw moment.

Central Moment of a PDF

Definitions of central moments

The nth central moment for a continuous probability distribution with density f(x) is defined as

$$\mu_n = \int_{-\infty}^{\infty} (x - \mu)^n f(x)\, dx$$

Observations

The most famous central moment is the second central moment, which is the variance.

It is easy to see that the first central moment is always 0.

Standardized Moment of a PDF

The nth standardized moment is defined by dividing the nth central moment by the nth power of the standard deviation:

$$\tilde{\mu}_n = \mu_n / \sigma^n$$

Observations

For the standardized moments, we have the following results:

1. The first standardized moment is always 0.
2. The second standardized moment is always 1.
3. The **third** standardized moment is the **skewness** of the distribution.
4. The **fourth** standardized moment is the **raw kurtosis** of the distribution.

Moments for discrete distributions

Similar definitions exist for discrete distributions.

Loosely speaking, you can replace the integrals by summations.

Technically, the moments are defined by using the notion of the expected value of a random variable.

For example:

Nth Raw Moment:	$E[X^n] = \sum_i x_i^n p_i$
Nth Central Moment:	$E[(X-\mu)^n] = \sum_i (x_i - \mu)^n p_i$

Higher order moments

Moments of the order higher than 4 are sometimes used in deep learning applications for example to extract shape based features.

Questions?