Econometrics

What is econometrics?

Arts and sciences of using statistics, mathematics, and economics theories in explaining empirical data.

Prerequisites:

Mathematics

- Linear Algebra
- Matrix Algebra

Statistics

- Probability Theories
- Estimation and Inference

Statistics

Statistics

Art and science of collecting, organizing, analyzing, interpreting, and presenting numerical data for the purpose of making a more effective decision.

Descriptive Statistics

Procedures used to organize and summarize masses of numerical data.

Inferential Statistics

Methods used to find out something about a population, based on a sample.

Descriptive Statistics

Methods of organizing, summarizing, and presenting data in an informative way, including:

- Frequency distribution
- Charts
- Measures of central tendency
  - Mean, Median, Mode
- Measures of dispersion
  - Range, Variance, Standard Deviation

Inferential Statistics

A decision, estimate, prediction, or generalization about a population, based on a sample.

Population

A collection of all possible individuals, objects, or measurements of interest.

Sample

A portion, or part, of the population of interest.

Sample must be a good representative of the population.

Types of Variables

Qualitative variables - nonnumeric

Gender, Types of Industry

Quantitative variables - numeric values

Discrete variables

Number of Male or Female in this class.

Continuous variables

Grade Point Average, Height, Weight, Stock price, SET Index.

Scales of Measurement

Measurement of data can be classified into 4 different levels

- Nominal Level
- Ordinal Level
- Interval Level
- Ratio Level

Descriptive Statistics

Measures of Central Tendency

Single value uses to represent a set of data.

Data are clustered close to it.

Measures of Dispersion

Variation in a set of data.

Can be used to evaluate the reliability of two or more averages.

Measures of Central Tendency

Arithmetic Mean

Sample Mean

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Population Mean

$$\mu = \frac{\sum_{i=1}^{N} X_i}{N}$$

Measures of Central Tendency

Properties of Arithmetic Mean

1. Every set of interval-level and ratio-level data has a mean.
2. All the data values are used in the calculation.
3. A set of data has only one mean.
4. The sum of the deviations from the mean equals 0.

$$\sum_{i=1}^{n}\left(X_i - \overline{X}\right) = 0$$

Measures of Central Tendency

Median:

The value in the middle set of the data.

Half the values are above it, half below.

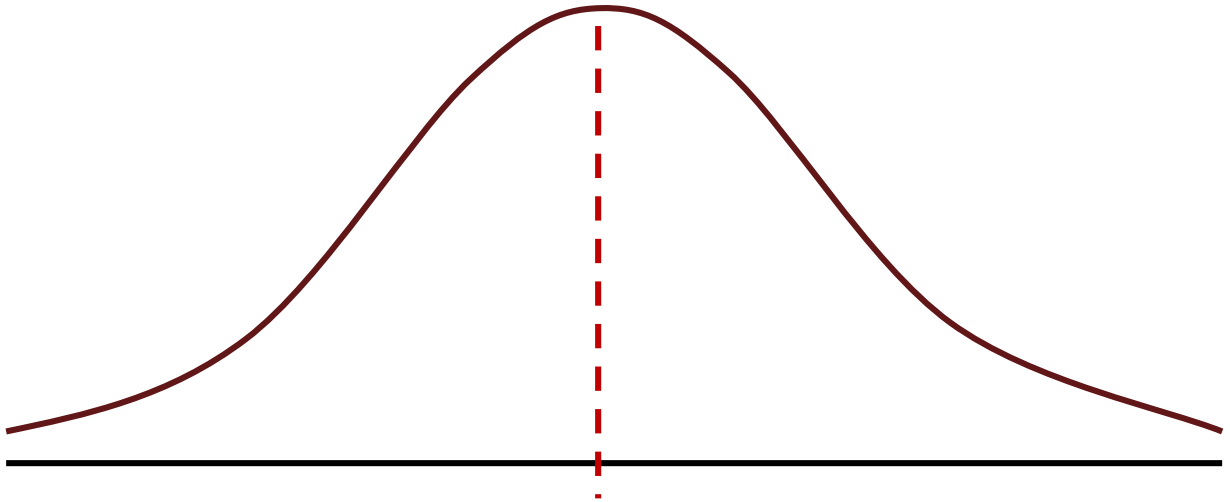
Mode:

The value that occurs most often in a set of data.

A set of data can have more than one mode.

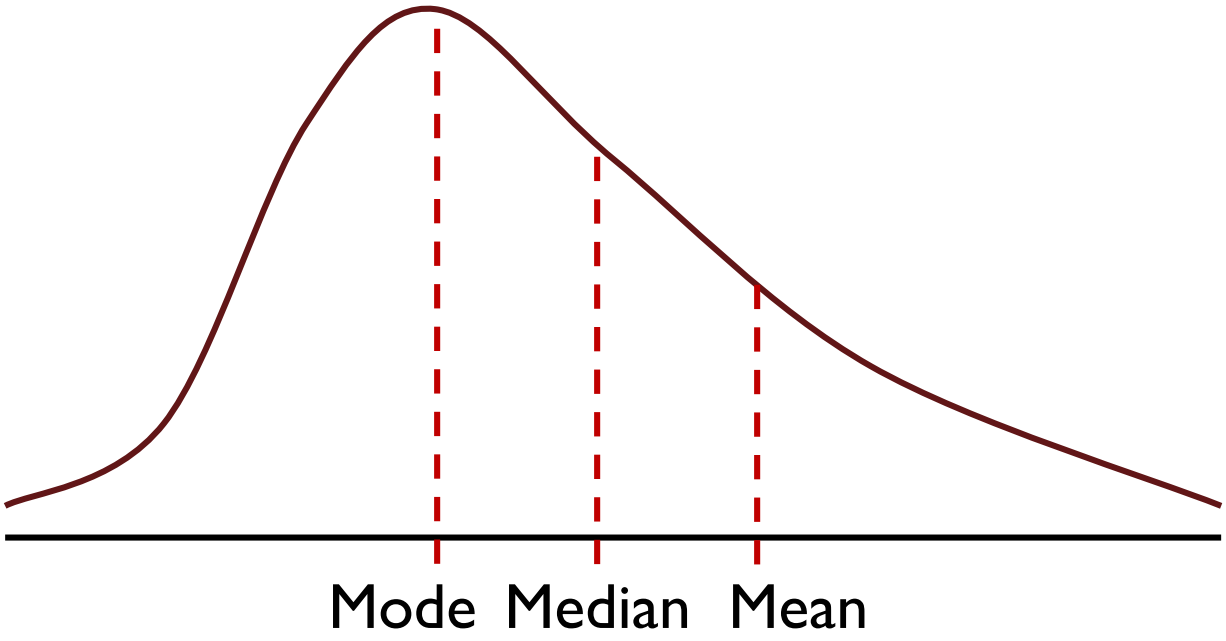
Measures of Central Tendency

For a symmetric distribution, mean, median, and mode are equal.



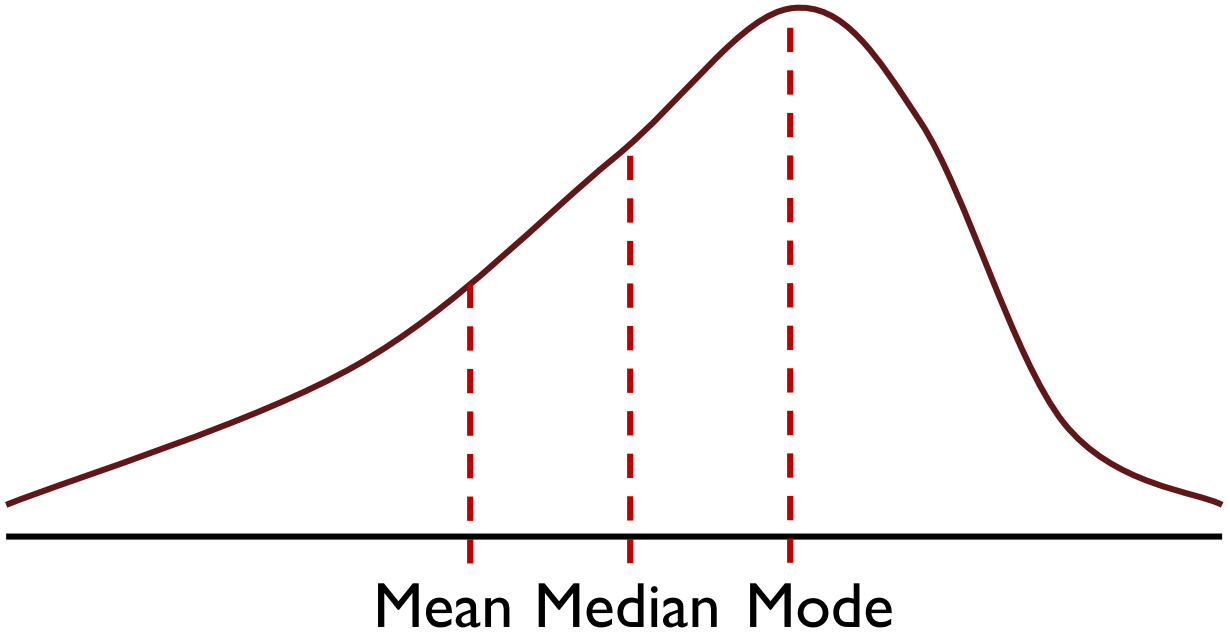
Measures of Central Tendency

Positively Skewed Distribution:



Measures of Central Tendency

Negatively Skewed Distribution:



Measures of Dispersion

Dispersion is the variation in a set of data.

Range:

Difference between the largest and the smallest observation in a set of data.

$$Range = Maximum\ value - Minimum\ value$$

Mean Absolute Deviation:

Sum of the absolute differences between each observation and the mean divided by number of observations.

$$MAD = \frac{\sum_{i=1}^{n} \left| X_i - \bar{X} \right|}{n}$$

Measures of Dispersion

Population Variance:

Mean of the squared deviations from the mean.

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$$

Sample Standard Deviation:

Square root of the variance.

$$S = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}}$$

Measures of Dispersion

Coefficient of Variation:

A measure of relative dispersion.

$$CV = \frac{s}{\bar{X}}(100)$$

Coefficient of Skewness:

A measure of Symmetry of a distribution.

Pearson's sk $$sk = \frac{3\left(\bar{X} - Median\right)}{s}$$

Software's sk $$sk = \frac{n}{(n-1)(n-2)}\left[\sum\left(\frac{X - \bar{X}}{s}\right)^3\right]$$

Descriptive Statistics

Key Summary Statistical Indices

- Number of Observation
- Mean
- Median
- Standard Deviation (s.d.)
- Minimum
- Maximum

To make sure that data (variables) have no extreme value.

Inferential Statistics

A decision, estimate, prediction, or generalization about a population, based on a sample.

Population: A collection of all possible individuals, objects, or measurements of interest.

Sample: A portion, or part, of the population of interest.

How well can sample represent population and conclusion on population drawn from sample?

Inferential Statistics

Key Issues in Inferential Statistics

- Whether Sample represents the population?
- Whether the results (findings) can be generalized?
  - Hypothesis Testing in order to confirm the generalization of the results (findings)
  - Probability Distribution of the data