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

Mathematics for Econometrics

Matrix and Vector

- Addition and Subtract of Matrix and Vector
- Multiplication of Matrix and Vector
- Inverse Matrix
- Identity Matrix
- Square Matrix
- Transpose
- Idempotent Matrix
- Determinant of Matrix
- Rank and Order of Matrix
- Singular and Nonsingular Matrix

Mathematics for Econometrics

Calculas

- Derivative & Differential
- Partial Derivative
- Total Differential
- Optimization
  - Maximization or Minimization
  - Constrained Optimization
  - First-order and Second-order Condition
  - Normal Equation

Mathematics for Econometrics

Rule of Differentiation

Constant-Function Rule:

$$y = f(x) = k$$

$$\frac{dy}{dx} = \frac{d}{dx}k = 0$$

Power-Function Rule:

$$y = f(x) = cx^n$$

$$\frac{dy}{dx} = \frac{d}{dx}cx^n = cnx^{n-1}$$

Mathematics for Econometrics

Rule of Differentiation Involving Two or More Functions of The Same Variable

Two functions of same variable x: $f(x)$ and $g(x)$

Sum-Difference Rule

$$\frac{d}{dx}\left[f(x) \pm g(x)\right] = \frac{d}{dx} f(x) \pm \frac{d}{dx} g(x)$$

Product Rule

$$\frac{d}{dx}\left[f(x)g(x)\right] = f(x)\frac{d}{dx} g(x) + g(x)\frac{d}{dx} f(x)$$

Mathematics for Econometrics

Rule of Differentiation Involving Two or More Functions of The Same Variable

Two functions of same variable x: $f(x)$ and $g(x)$

Quotient Rule

$$\frac{d}{dx}\frac{f(x)}{g(x)} = \frac{g(x)f'(x) - f(x)g'(x)}{g^2(x)}$$

Chain Rule

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx} = f'(y)g'(x)$$

$$\frac{dz}{dw} = \frac{dz}{dy}\frac{dy}{dx}\frac{dx}{dw} = f'(y)g'(x)h'(w)$$

Mathematics for Econometrics

Rule of Differentiation Involving Functions of Different Variables

Chain Rule

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx} = f'(y)g'(x)$$

$$\frac{dz}{dw} = \frac{dz}{dy}\frac{dy}{dx}\frac{dx}{dw} = f'(y)g'(x)h'(w)$$

Mathematics for Econometrics

Rule of Differentiation

Exponential Function :

$$\frac{d}{dx}e^x = e^x$$

$$\frac{d}{dx}ke^{nx} = kne^x$$

Logarithmic Function:

$$\frac{d}{dx}\ln x = \frac{1}{x}$$

$$\frac{d}{dx}k\ln x^n = \frac{kn}{x}$$

Basic Statistics for Econometrics

Descriptive Statistics

Measures of Central Tendency

e.g. Mean, Median, Mode

Measures of Dispersion

e.g. Variance, Standard Deviation, Covariance

Basic Statistics for Econometrics

Inferential Statistics

Probability Theory

- Probability Concept
- Random Variable
- Probability Distribution
- Probability Density Function
- Cumulative Distribution Function

Expectation of Random Variable

Variance of Random Variable

Basic Statistics for Econometrics

Probability Distribution

Discrete Distribution

- Binomial Distribution
- Poisson Distribution

Continuous Distribution

- Normal Distribution
- t Distribution
- Chi-square Distribution
- F Distribution

Basic Statistics for Econometrics

Hypotheses Testing

- Null Hypothesis vs Alternate Hypothesis
- Sampling Distribution
- Sampling Error
- Point & Interval Estimation
- Confidence Interval
- One-tail & Two-tail Test
- Z-test
- t-test
- Analysis of Variance (ANOVA) – F-test
- χ^2-test

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.

Examples

Population of all 5xx Listed Companies in SET

100 Companies

Sample of 100 companies of the SET

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

Scales of Measurement

Nominal Level

Data that is classified into categories and cannot be arranged in any particular order.

Nominal level variables must be:

Mutually exclusive

An individual, object, or measurement is included in only one category.

Exhaustive

Each individual, object, or measurement must appear in one of the categories.

Scales of Measurement

Ordinal Level

involves data arranged in some order, but the differences between data values cannot be determined or are meaningless.

Interval Level

Similar to the ordinal level, with the additional property that meaningful amounts of differences between data values can be determined. There is no natural zero point. Interval scale distance between values is constant.

Scales of Measurement

Ratio Level

The interval level with an inherent zero starting point. Differences and ratios are meaningful for this level of measurement.

Data have a meaningful zero point.

Summarizing Data

Frequency Distribution

A grouping of data into categories showing the number of observations in each mutually exclusive category.

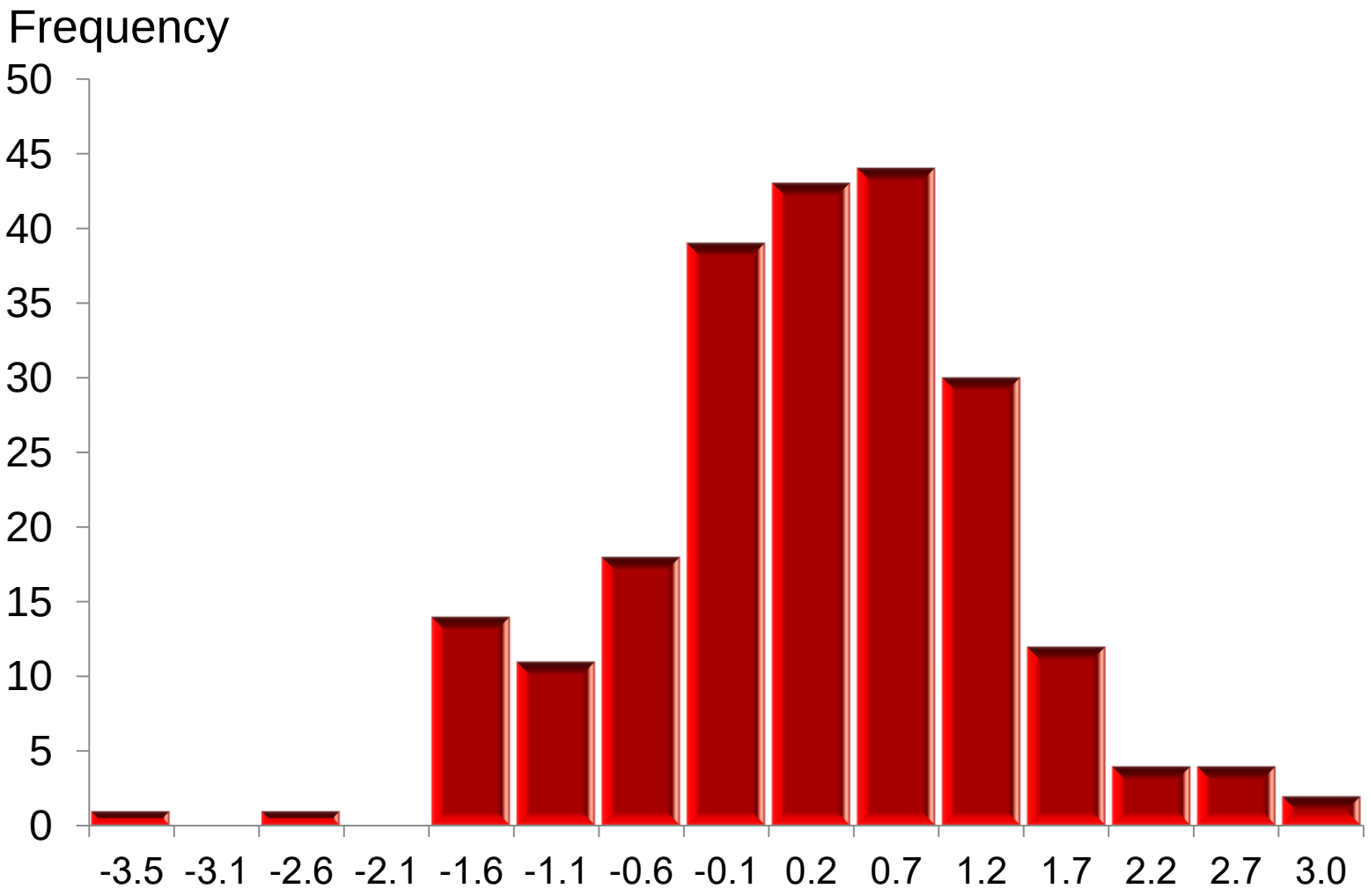
Histogram portrays the number of frequencies in each class in the form of bars.

Frequency Polygon consists of line segments connecting the points formed by the intersections of the class midpoints and the class frequencies.

Example – Frequency Table

Return	Mid	Frequency	Percent
Less than -3.3	-3.5	1	0.45%
-3.3 to -2.9	-3.1	0	0.00%
-2.8 to -2.4	-2.6	1	0.45%
-2.3 to -1.9	-2.1	0	0.00%
-1.8 to -1.4	-1.6	14	6.28%
-1.3 to -0.9	-1.1	11	4.93%
-0.8 to -0.4	-0.6	18	8.07%
-0.3 to 0.0	-0.1	39	17.49%
0.1 to 0.4	0.2	43	19.28%
0.5 to 0.9	0.7	44	19.73%
1.0 to 1.4	1.2	30	13.45%
1.5 to 1.9	1.7	12	5.38%
2.0 to 2.4	2.2	4	1.79%
2.5 to 2.9	2.7	4	1.79%
More than 3.0	3.0	2	0.90%
Total		223	100.00%

Example - Histogram



Graphic Presentation

Line Charts

Bar Charts

- Horizontal bar charts
- Vertical bar charts

Stacked Bar Charts

Pie Charts

Statistics for Financial Analysis

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

Weighted Mean:

$$\bar{X}_w = w_1X_1 + w_2X_2 + w_3X_3 + ... + w_nX_n$$

$$where: w_1 + w_2 + w_3 + ... + w_n = 1$$

Geometric Mean:

$$GM = \sqrt[n]{(X_1)(X_2)...(X_n)}$$

Rate of Change from one period to another:

$$GM = \sqrt[n-1]{\frac{Value\ at\ end\ of\ period}{Value\ at\ start\ of\ period}} - 1$$

Measures of Central Tendency

Median:

The value in the middle set of the data.

Half the values are above it, half below.

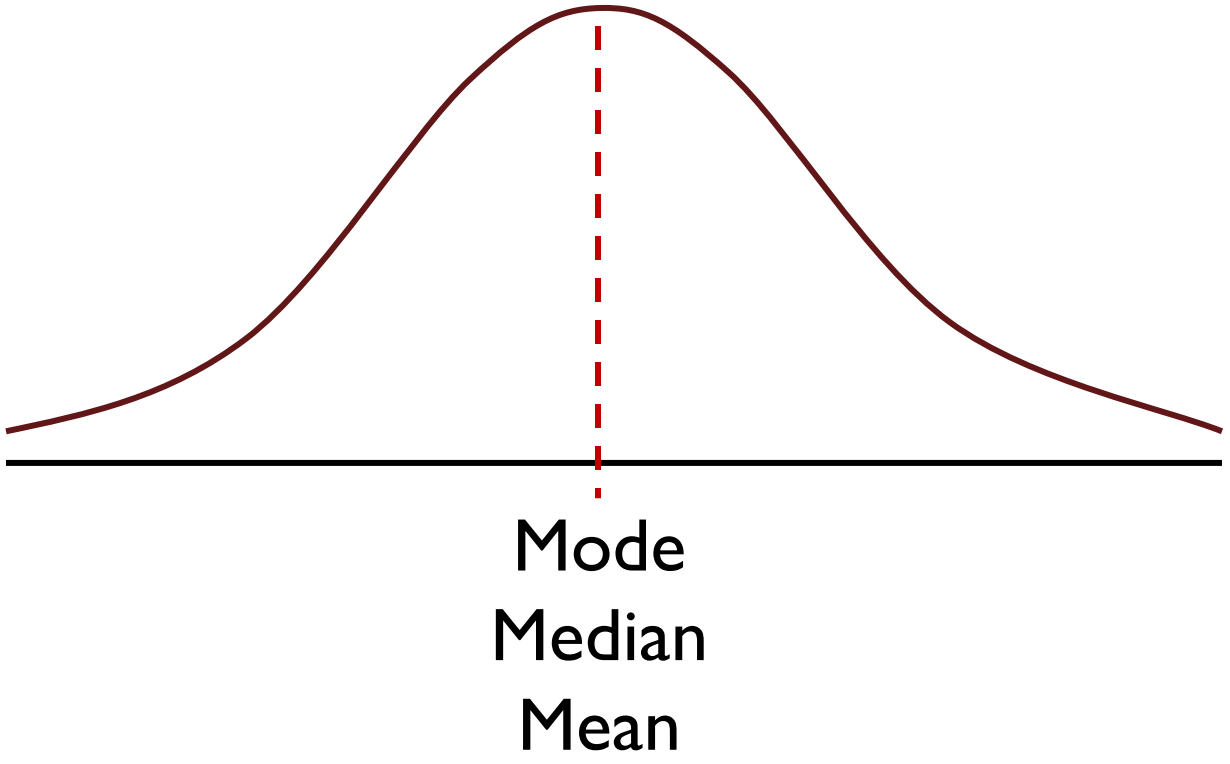
Mode:

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

A set of data can have more than one mode.

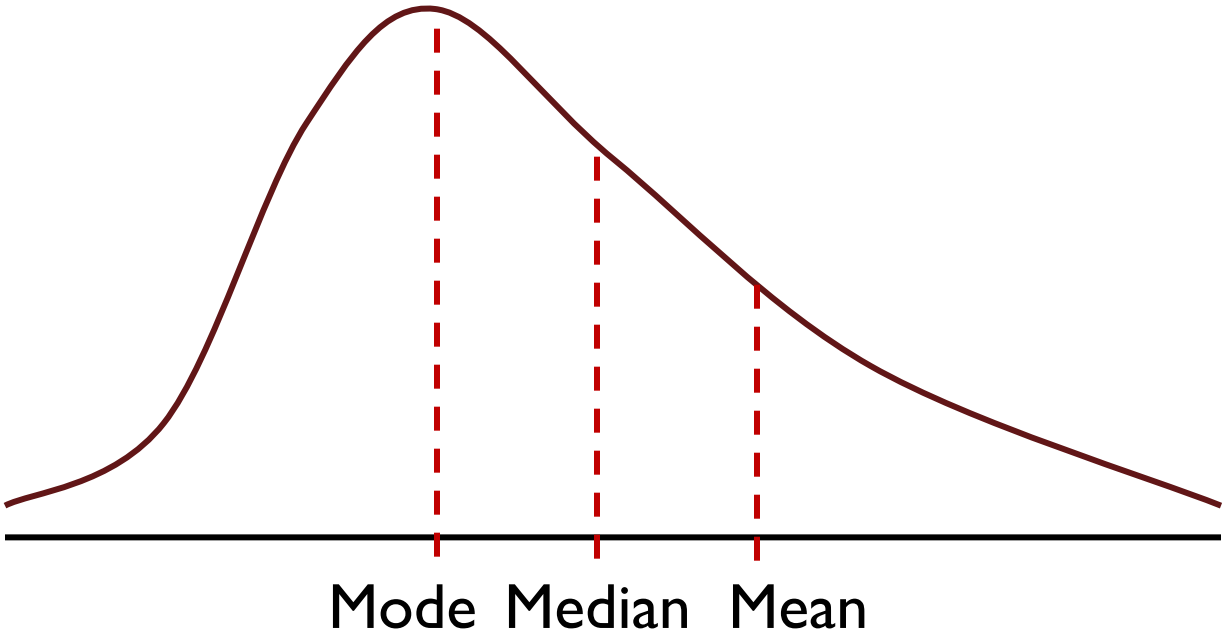
Measures of Central Tendency

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



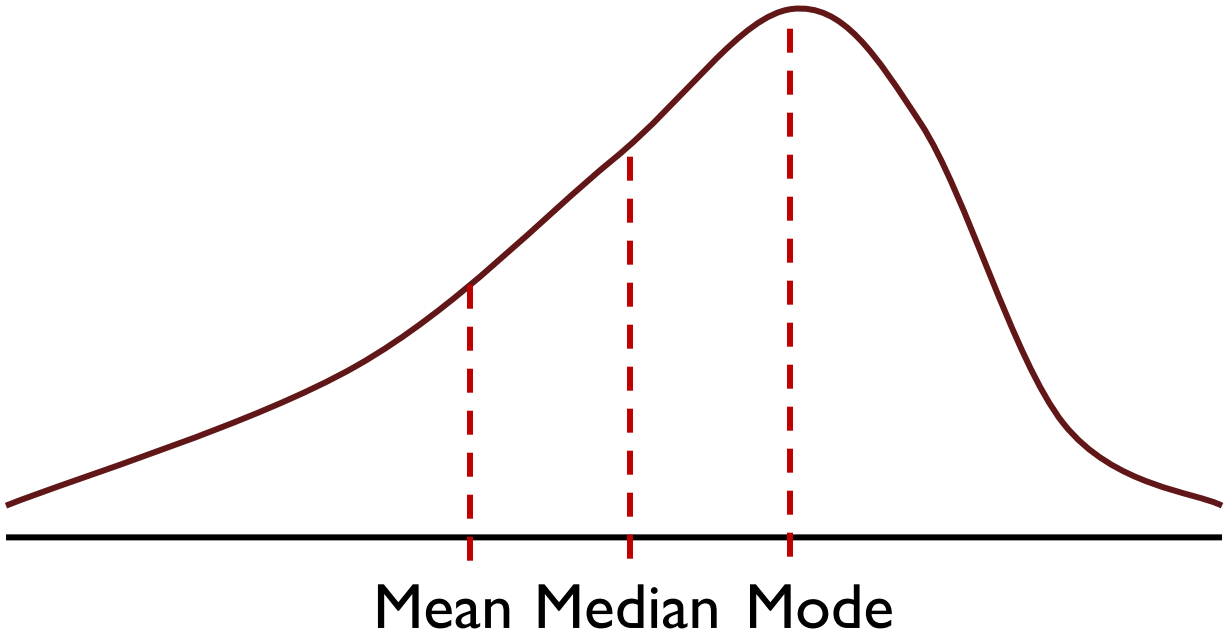
Measures of Central Tendency

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.

$$Range = Highest\ value - Lowest\ value$$

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