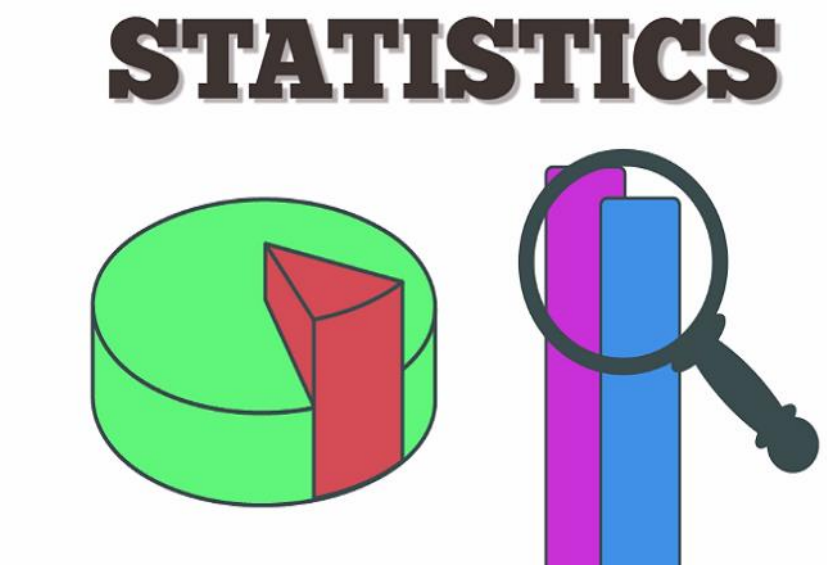
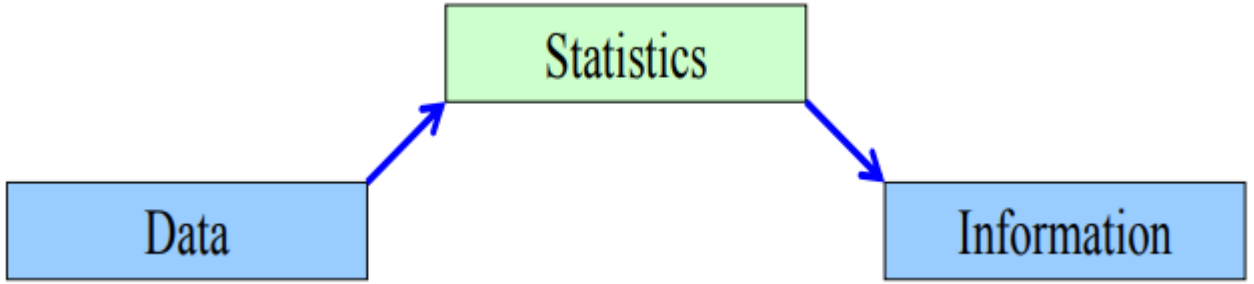
Introduction to Statistics and Analytics



What is Statistics?

- “Statistics is a way to get information from data.”



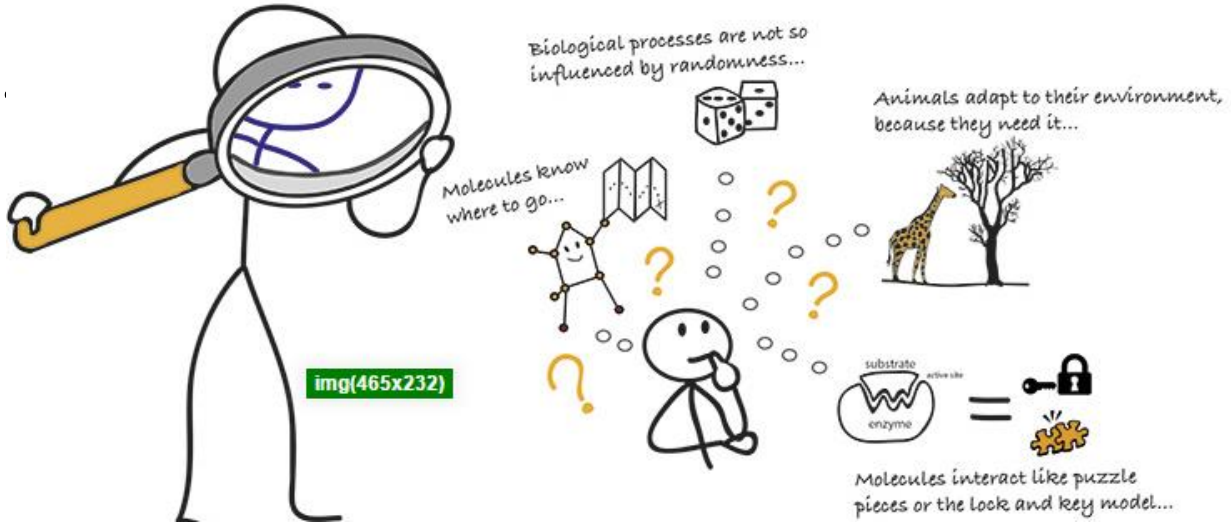
Data: Facts, especially numerical facts, collected together for reference or information.

Information: Knowledge communicated concerning some particular fact.

- Statistics is a tool for creating an understanding from a set of numbers.

Key Statistical Concepts

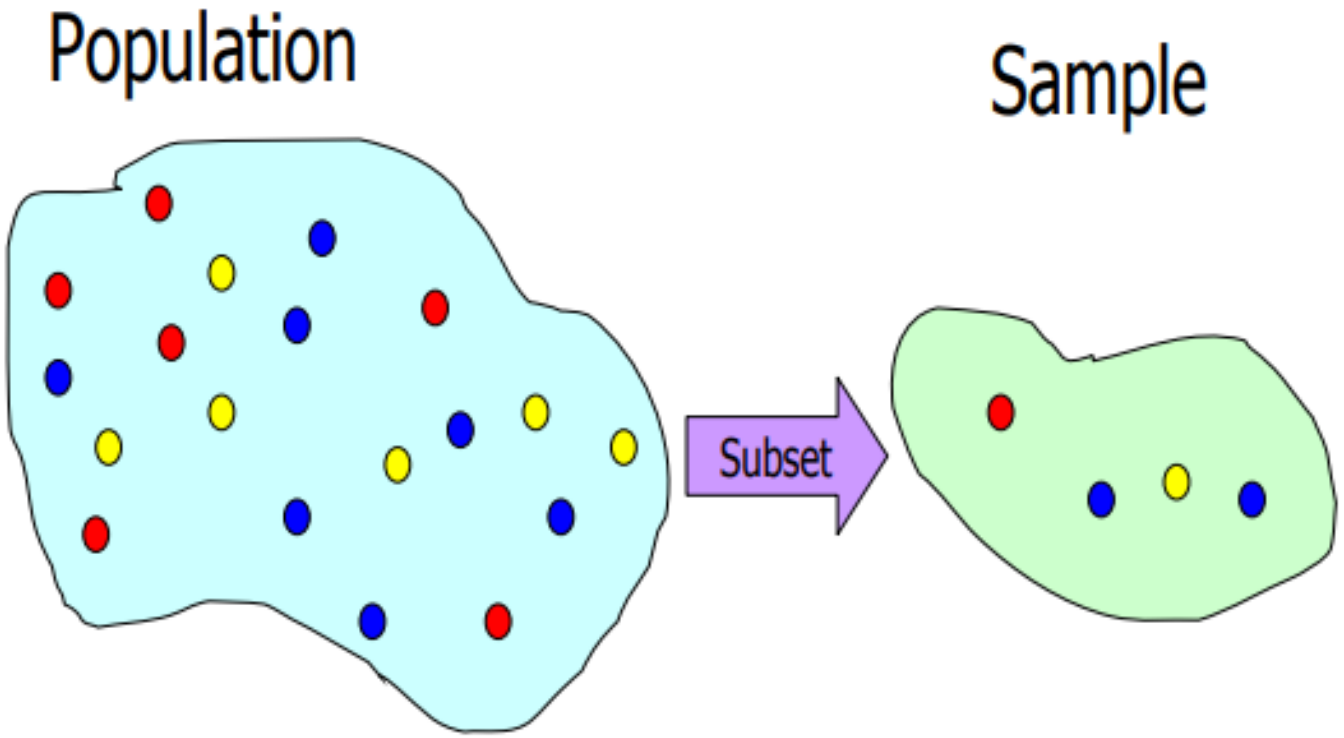
- Population
  - The population is all the individuals in a group.
  - frequently very large; sometimes infinite.
- Sample
  - The sample is some of the individuals in a group.
  - Potentially very large, but less than the population.
- Parameter
  - Defines a characteristic of the population
- Statistic
  - Defines a characteristic of the sample.



Key Statistical Concepts

- Independent Variable.
  - The variable that experimenters choose to manipulate; it is usually plotted along the x-axis of a graph.
- Dependent Variable.
  - The dependent variable of a study is the variable that experimenters choose to measure during an experiment; it is usually plotted along the y-axis of a graph.

Population vs Sample



Populations have Parameters,

Samples have Statistics.

Descriptive Statistics

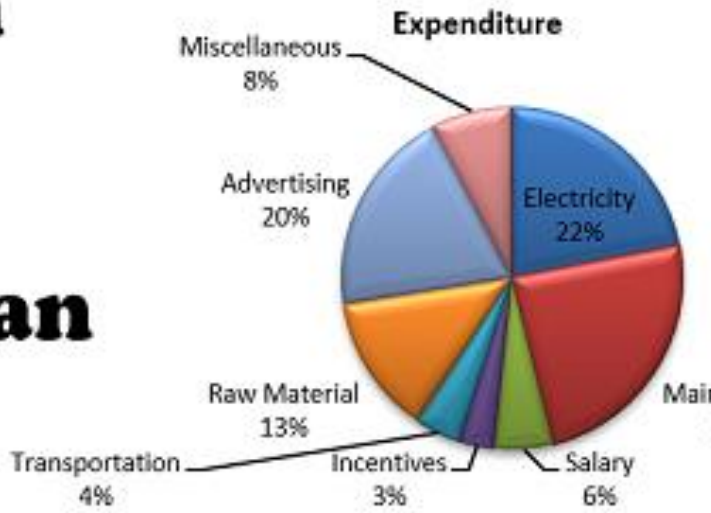
- Are methods of organizing, summarizing, and presenting data in a convenient and informative way.
  - Graphical Techniques
  - Numerical Techniques
- The actual method used depends on what information we would like to extract.
  - measure(s) of central location? and/or
    - Mean/ Median/ Mode
  - measure(s) of variability (dispersion)?
    - Range/ IQR/ Variance/ Standard Deviation

Descriptive Statistics

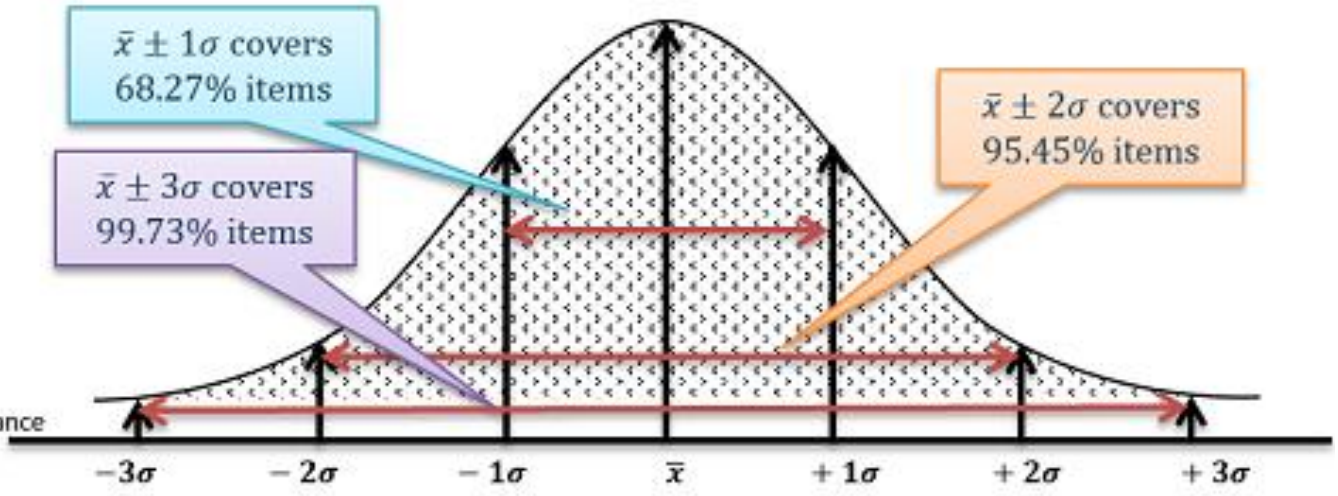
Mean

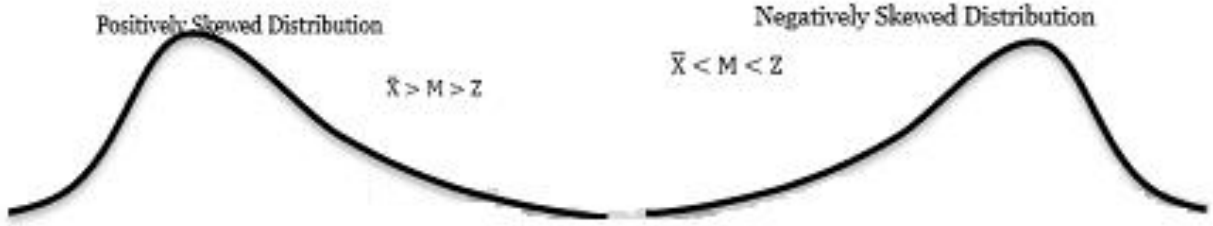
Median

Mode



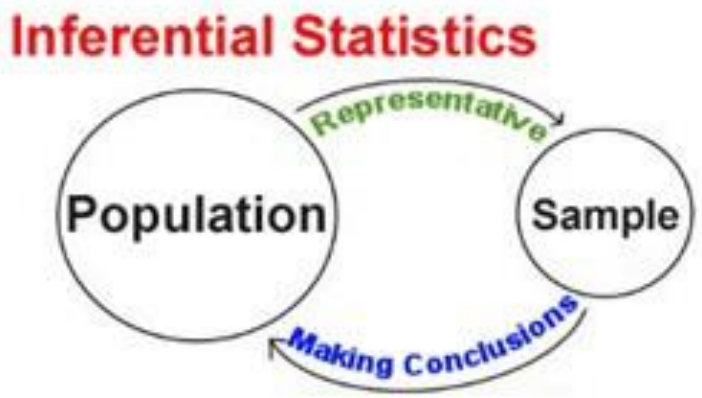
$$Std.Dev.\sigma = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}}$$



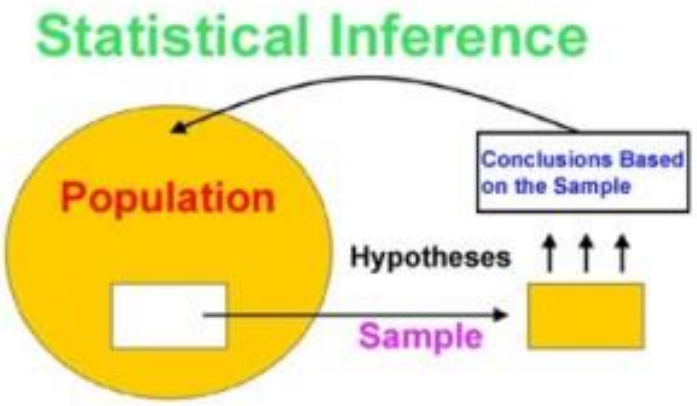


Inferential Statistics

- Descriptive Statistics describe the data set that's being analyzed.
- But doesn't allow us to draw any conclusions about the data.
- Hence we need another branch of statistics: inferential statistics.



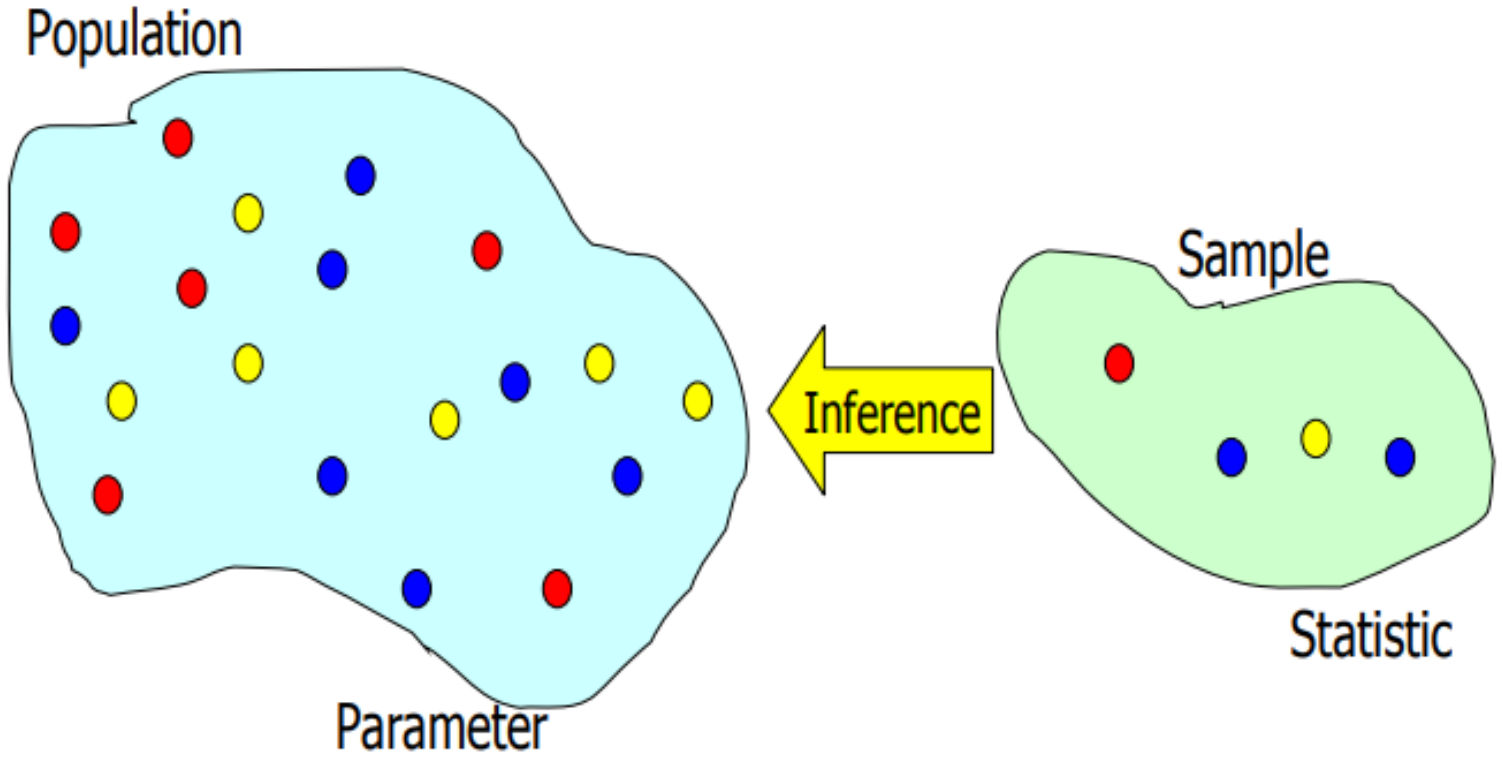
- Inferential statistics is also a set of methods,
- But it is used to draw conclusions or inferences about
  - characteristics of populations based on data from a sample.



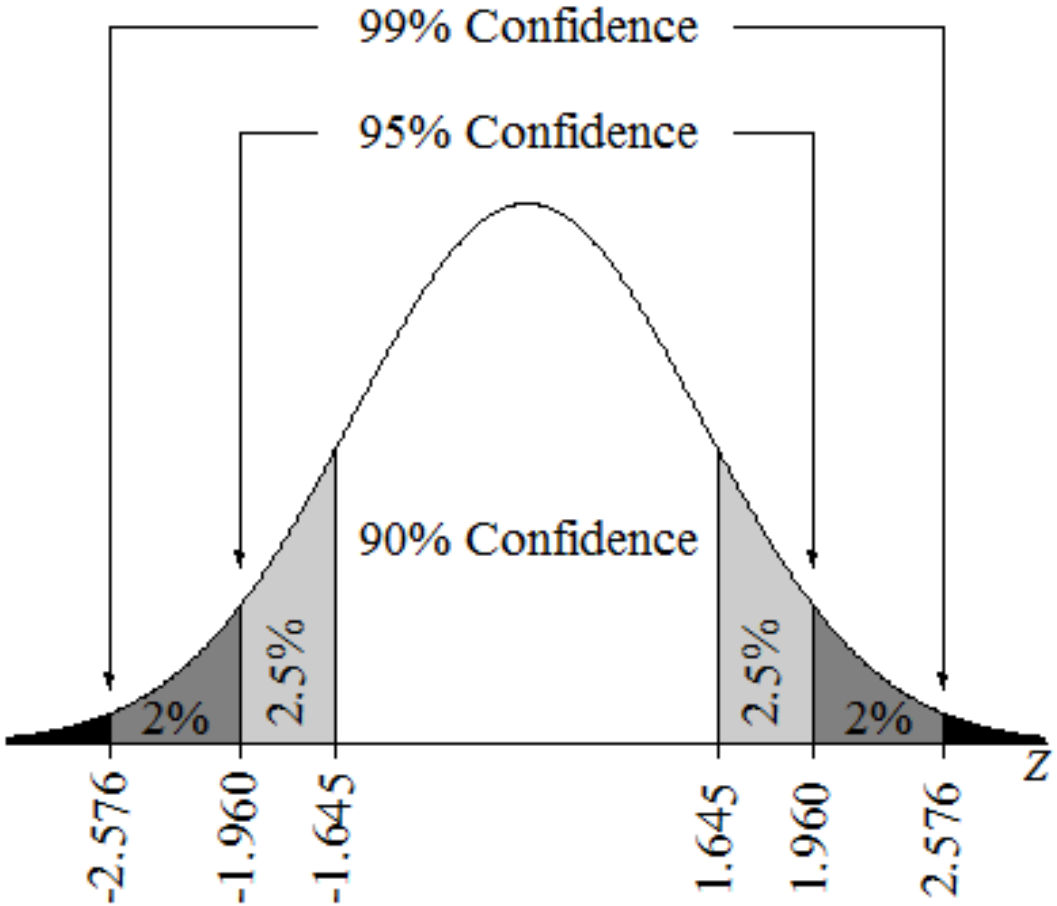
Statistical Inference. . .

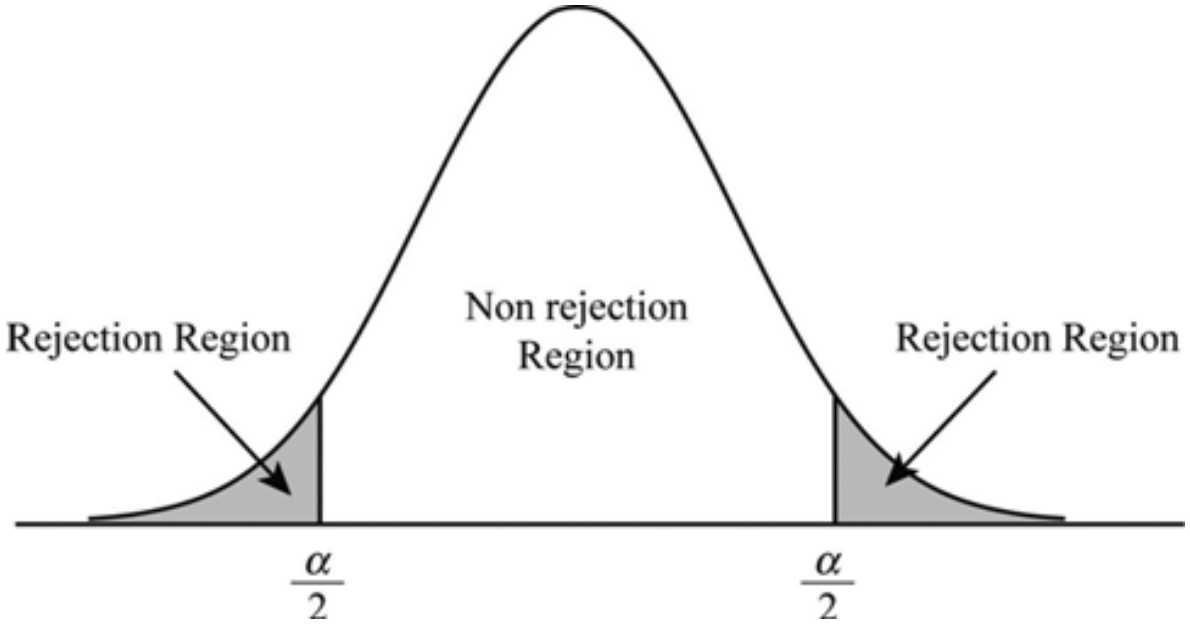
- Statistical inference is the process of making an
  - estimate,
  - prediction, or
  - decision
- about a population based on a sample.

We use statistics to make inferences about parameters.



Confidence Levels





Thanks