# Unit 1: Introduction

Statistics 571: Statistical Methods

Ramón V. León

# What is Statistics?

- Statistics
  - Science of collecting and analyzing data for the purpose of drawing conclusions and making decisions
  - Provides data collection methods to reduce biases, and analysis methods to identify patterns and draw inference from noisy data
- Statistical tasks
  - Collecting data
  - Summarizing and exploring data
  - Drawing conclusions and making decisions based on data

# Statistical Concepts and Terms

- Observational studies such as surveys
- Experimental studies
- Descriptive statistics or exploratory data analysis
  - Summarizes data
- Inferential statistics or confirmatory data analysis
  - Fitting models to data
  - Drawing conclusions and making decisions
- A *populati*on is a collection of all units of interest
- A *sample* is a subset of a population that is actually observed
- A *variable* is a measurable property or attribute associated with each unit of a population

# Statistical Concepts and Terms

- A *parameter* is a numerical characteristic of a population defined for each variable of interest
- A *statistic* is a numerical characteristic of a sample defined for each variable of interest
- Statistics are used to infer the values of parameters
- A *random sample* gives an equal chance to every groups of unit of the population to enter the sample
  - One form of unbiased sample
- In probability we assume that the population and its parameters are known and compute the probability of drawing a particular sample
- In statistics we assume that population parameters are unknown and use a sample to infer their value

# Statistical Concepts and Terms

- Different samples give different estimates of population parameters (called *sampling variability*)
- Sampling variability leads to sampling error
- One goal of statistic is to quantify sampling error
- Probability is deductive
- Statistics is inductive

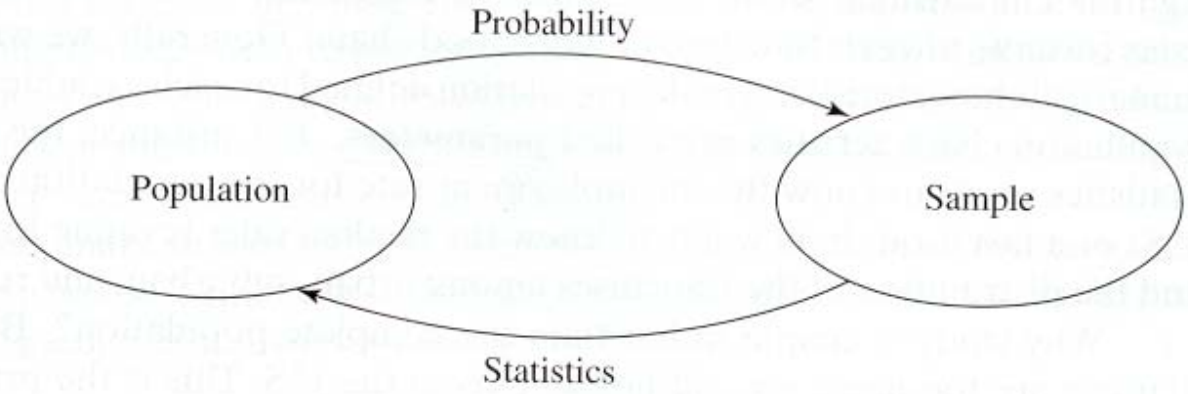


**Figure 1.1** Relationship between Probability and Statistics*