Chapter 1
Introduction to Statistics

What is Statistics?

Statistics is the science of learning from data.

Statistics: Statistics is the science of collecting, analyzing, presenting, and interpreting data, as well as of making decisions based on such analyses.

Educated Guess vs. Pure Guess: Every day we make decisions that may be personal, business related, or of some other kind. Usually these decisions are made under conditions of uncertainty. Many times, the situations or problems we face in the real world have no precise or definite solution. Statistical methods help us make scientific and intelligent decisions in such situations. **Decisions made by using statistical methods are called educated guesses. Decisions made without using statistical (or scientific) methods are pure guesses** and, hence, may prove to be unreliable.

Basic Terms

Element or Member: An element or member is a specific subject or object (for example, a person, firm, item, state, or country) about which the information is collected.

Population: A population consists of all elements—individuals, items, or objects—whose characteristics are being studied. The population that is being studied is also called the target population.

Sample: A portion of the population selected for study is referred to as a sample.

Variable: A variable is a characteristic under study that assumes different values for different elements. In contrast to a variable, the value of a constant is fixed.

Observation or Measurement: The value of a variable for an element is called an observation or measurement.

Data Set: A data set is a collection of observations on one or more variables.

Survey: The collection of information from the elements of a population or a sample is called a survey.

Census: Survey that includes every member of the population is called a census.

Sample Survey: Survey that includes only a portion of the population is called a sample survey.

Representative Sample: A sample that represents the characteristics of the population as closely as possible is called a representative sample.

Types of Statistics

Descriptive Statistics (Chapters 2, 3): Descriptive statistics consists of methods for organizing, displaying, and describing data by using tables, graphs, and summary measures.

Inferential Statistics (Parts of Chapter 7 and Chapters 8, 9) : Inferential statistics consists of methods that use sample results to help make decisions or predictions about a population.

A major portion of statistics deals with making decisions, inferences, predictions, and forecasts about populations based on results obtained from samples. For example, we may make some decisions about the political views of all college and university students based on the political

Probability (Chapters 4-6, parts of 7): Probability, which gives a measurement of the likelihood that a certain outcome will occur, acts as a link between descriptive and inferential statistics. Probability is used to make statements about the occurrence or nonoccurrence of an event under uncertain conditions.

Random vs. Non Random Sample

Random Sample: A sample drawn in such a way that each element of the population has a chance of being selected is called a random sample. If all samples of the same size selected from a population have the same chance of being selected, we call it simple random sampling. Such a sample is called a simple random sample.

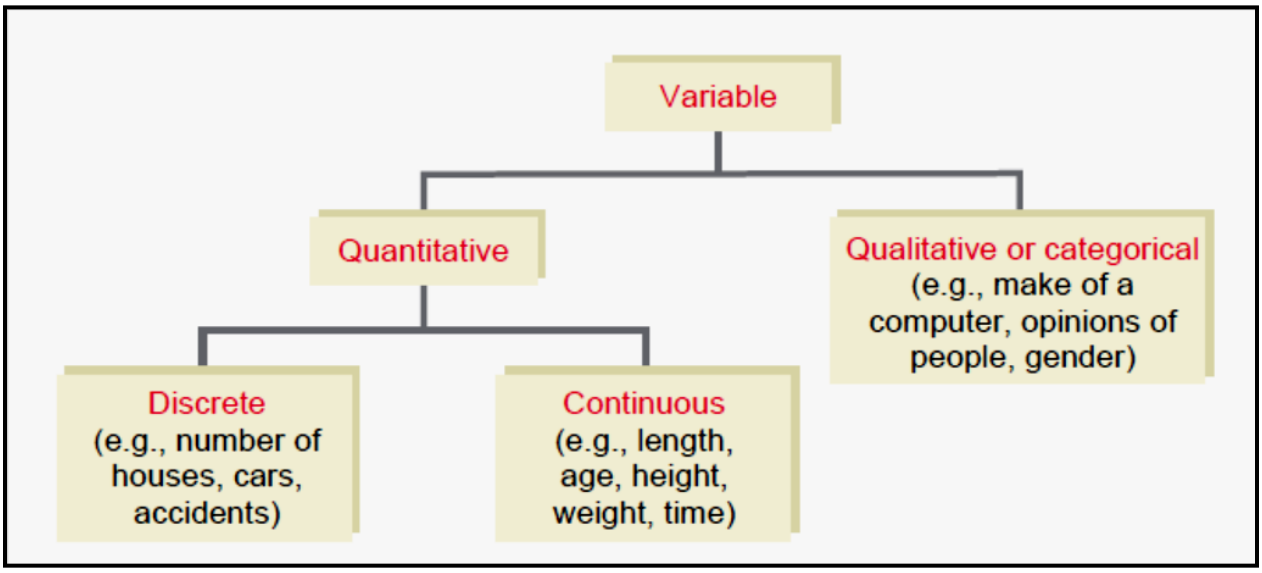
Non Random Sample: A sample drawn in such a way that each element of the population does not have a chance of being selected is called a random sample.

Sampling with Replacement vs. Sampling without Replacement

Sampling with Replacement: In sampling with replacement, each time we select an element from the population, we put it back in the population before we select the next element. Thus, in sampling with replacement, the population contains the same number of items each time a selection is made. As a result, we may select the same item more than once in such a sample.

Sampling without Replacement: Sampling without replacement occurs when the selected element is not replaced in the population. In this case, each time we select an item, the size of the population is reduced by one element. Thus, we cannot select the same item more than once in this type of sampling. Most of the time, samples taken in statistics are without replacement.

Types of Variables



Quantitative Variable: A variable that can be measured numerically is called a quantitative variable. The data collected on a quantitative variable are called quantitative data.

Discrete Variable: A variable whose values are countable is called a discrete variable. In other words, a discrete variable can assume only certain values with no intermediate values.

Continuous Variable: A variable that can assume any numerical value over a certain interval or intervals is called a continuous variable.

Qualitative or Categorical Variable: A variable that cannot assume a numerical value but can be classified into two or more nonnumeric categories is called a qualitative or categorical variable. The data collected on such a variable are called qualitative data.

Types of Data

Cross-Section Data: Data collected on different elements at the same point in time or for the same period of time are called cross-section data.

Time-Series Data: Data collected on the same element for the same variable at different points in time or for different periods of time are called time-series data.

Types of Statistics

Descriptive Statistics (Chapters 2, 3): Descriptive statistics consists of methods for organizing, displaying, and describing data by using tables, graphs, and summary measures.

Inferential Statistics (Parts of Chapter 7 and Chapters 8, 9) : Inferential statistics consists of methods that use sample results to help make decisions or predictions about a population.

A major portion of statistics deals with making decisions, inferences, predictions, and forecasts about populations based on results obtained from samples. For example, we may make some decisions about the political views of all college and university students based on the political

Probability (Chapters 4-6, parts of 7): Probability, which gives a measurement of the likelihood that a certain outcome will occur, acts as a link between descriptive and inferential statistics. Probability is used to make statements about the occurrence or nonoccurrence of an event under uncertain conditions.

Problems from Chapter 1 (from 8th Edition of the book)

Solved in class: 1.15, 1.22, 1.27, 1.31

Practice at home: 1.10, 1.12, 1.14, 1.28, 1.34