Statistics Chapter 1

STATISTICS is the science of conducting studies to collect, organize, summarize, analyze, and draw conclusions from data.

Section 1-2

Statisticians collect information about variables which describe events.

A VARIABLE is a characteristic that can assume different values.

DATA are the values that the variables can assume.

The values of RANDOM VARIABLES are determined by chance.

Statistics can be divided into two main branches depending on how the data is used.

1) DESCRIPTIVE statistics consists of the collection, organization, summarization and presentation of data.

2) INFERENTIAL statistics involves making inferences from samples to populations. This involves PROBABILITY: the chance of an event occurring. Performing estimations, hypothesis testing, determining relationships among variables and making predictions all fall in this category.

A POPULATION consists of all subjects that are being studied.

A SAMPLE is a group of subjects selected from a population.
Section 1-3 Variables and Types of Data

Types of variables:

QUALITATIVE variables can be placed into distinct categories.

QUANTITATIVE variables are numerical and can be put in order. They are either discrete or continuous.

DISCRETE variables are countable.

CONTINUOUS variables can assume an infinite number of values between any two specific values, they are obtained by measuring.

Measurement scales describe how variables are categorized, counted, or measured.

NOMINAL level – there are mutually exclusive, exhaustive categories, no order or ranking.

ORDINAL level – categories can be ranked, but precise differences between ranks do not exist.

INTERVAL level – data is ranked, and precise differences between units exist. There is no meaningful zero. Negative values do not make you laugh.

RATIO level – like interval, but there is a true zero and ratios exist.
Section 1-4 Basic Sampling Techniques

RANDOM – Use chance methods or random numbers. Each subject has an equal chance of being selected.

SYSTEMATIC – Select every nth subject.

STRATIFIED – Divide the population into groups (strata) based on a characteristic important to the study, then sample from each group.

CLUSTER – Use an intact group that represents the population.

There are many other sampling techniques, including the CONVENIENCE sample.

Section 1-5 Observational and Experimental Studies

In an OBSERVATIONAL STUDY the researcher observes what is happening and tries to draw conclusions from the observations.

In an EXPERIMENTAL STUDY the researcher manipulates one variable and tries to determine how it influences other variables.

The INDEPENDENT or EXPLANATORY variable is the one that is manipulated.

The DEPENDENT or OUTCOME variable is the one studied to see if it changes due to the manipulated variable.

In an experimental study the subjects are usually placed into two groups. The group that is manipulated is the TREATMENT group, the group that does not receive any special treatment is called the CONTROL group. If the groups are not chosen randomly, the study is quasi-experimental. If the subjects know they are participating in a study, it may have an effect on the dependent variable. This is known as the HAWTHORNE EFFECT.
Section 1-6 Uses and Misuses of Statistics

Statistics are used to inform, but also to influence the public, so be critical of:

Suspect Samples – is the sample large enough, does it represent the population

Ambiguous Averages – is the type of average (mean, median, mode) identified

Changing the Subject – percentages and actual values can have different impacts

Detached Statistics – incomplete comparisons, like “50% faster”

Implied Connections – watch for phrases like “may help”, “studies suggest”, “in some people”

Misleading Graphs – visual representations can be distorted

Faulty Survey Questions – the phrasing of questions can influence the way people answer them