

Lesson 13.1: Basic Principles of Statistics

Statistics is the **science** of **data**.

- **Individuals** are the objects (**people**, **animal**, or **things**) described by a set of data.
- A **variable** is any **characteristic** of an individual.
- A **categorical** variable places an individual into one of several **groups** or **categories**.
- A **quantitative** variable takes **numerical** values for which it **makes sense** to find an **average**.

Identify each variable as categorical or quantitative.

Categorical

type of wood

type of water repellent

paint color

Quantitative

paint thickness

weathering time

The population in a statistical study is the entire group of individuals we want information about. A census collects data from every individual in the population. A sample is a subset of individuals in the population from which we actually collect data.

- A parameter is a number that describes some characteristic of the population.
- A statistic is a number that describes some characteristic of a sample.

Circle the Statistic and Underline the Parameter

A telemarketing firm in a large city uses a device that dials residential telephone numbers in that city at random. Of the first 100 numbers dialed 48% are unlisted. This is not surprising because 52% of all residential phones in the city are unlisted.

A random sample of female college students has a mean height of 64.5 inches, which is greater than the 63 inch mean height of all adult American women.

You can hardly go a day without hearing the results of a statistical study. Can we trust the results? You will learn that it depends on how the data were produced. It also becomes important to know where the data came from.

Types of Studies

- An **Observational Study** observes individuals and **measures** variables of **interest** but does **not** attempt to **influence** the responses.
 - A **sample survey** uses an **organized plan** to choose a sample that **represents** some specific population, and then bases **conclusions** about the population on **data** from the sample.
- An **experiment** deliberately **imposes** some **treatment** on individuals to **measure** their responses.

The method of the selection of a sample is a key element to successful observational studies. It is important that a sample truly represents the populations from which it is taken. For this reason, it is important to know how to avoid sampling poorly and how to sample well.

How to Sample Poorly

The design of a statistical study shows bias if it would consistently underestimate or consistently overestimate the value you want to know.

- Choosing individuals from the population who are easy to reach results in a convenience sample.
- A voluntary response sample consists of people who choose themselves by responding to a general invitation.

Both of these sampling methods suffer from bias due to personal choice.

How to Sample Well

- The best way to avoid problems of bias is to let chance choose the sample. Random Sampling involves using a chance process to determine which members of a population are included in the sample.
- The most basic method of random sampling is a simple random sample (SRS). An SRS of size n is chosen in such a way that every group of n individuals in the population has an equal chance to be selected as the sample. Examples include drawing names from a hat, using dice or other random phenomenon, using a random number table, or using random number generators.

Random Assignment Experimentation

- A specific **condition** applied to **individuals** in an experiment is called a **treatment**. Ethically, a random sample **cannot** always be selected for **experiments** because **treatments** are imposed. Therefore to **reduce** bias in experiments, it is necessary to randomize in a different way than **selecting** participants. In an experiment, **random assignment** means that **experimental units** are assigned to **treatments** using a chance process.

Example

A firm wants to understand the attitudes of its minority managers toward its system for assessing management performance. It uses a computer program to randomly select 10 managers from a list of all their minority managers.

- Does this sampling method create a representative group? Why or Why not?

Yes. They were randomly selected

- Will the sample proportion of the minority managers positive attitude underestimate, overestimate, or be approximately the same as the actual population proportion? Explain your answer.

↓ if it's representative

Approximately the same because we have a representative group.

Example

While surfing the internet a pop-up window, asks you to complete a survey about how much time you spend on the internet per week.

- Does this sampling method create a representative group? Why or Why not?

No. It's only asking those online and they volunteer to complete the survey

- Will the sample mean of hours spent on the internet underestimate, overestimate, or be ~~approximately the same~~ as the actual population mean? Explain your answer.

Overestimate. They have to be on the internet to get the pop-up, so they spend a lot of time online.

Example

Researchers are interested to determine whether storing batteries in a freezer make them last longer. To find out, they randomly select 100 AA batteries from a battery warehouse. They then randomly assigned 50 to be stored in a freezer and the other 50 stored at room temperature for 3 years. At the end of the time period, the battery charges will be tested.

- What type of study are they conducting? Explain.

Experimental. Treatments imposed: freezer vs room temp.

- How is randomization used in the design of the study?

① Randomly selected sample

② Randomly assigned to groups

Example

A psychologist wants to study the effects of failure and frustration on the relationships among members of a work team.

- What type of study should she conduct? Explain.

Observational. Survey: they will tell us about their work experience.

- How might randomization be used in the design of the study you have selected?

Randomly select sample from several different work environments