

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.

Example: How can we help wood surfaces resist weathering, especially when restoring historic wooden buildings? In a study of this question, researchers prepared wooden panels and then exposed them to weather. Here are some of the variables recorded: type of wood (yellow poplar, pine, cedar); type of water repellent (solvent-based, water-based); paint thickness (millimeters); paint color (white, gray, light blue); weathering time (months). Identify each variable as categorical or quantitative.

type of wood
categorical

type of water repellent
categorical

paint thickness
quantitative

paint color
categorical

weathering time
quantitative

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.

For each bold number in the following descriptions, circle it if it is a statistic and underline it if it is a parameter.

Example: 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.

Example: 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 in Experiments

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.

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

Yes, they used random assignment which reduces bias.

- b. 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.

It will most likely be approximately the same since it was a random sample taken.

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.

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

No, it's only asking those on the internet and they volunteer to complete it.

- b. 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, since they are obviously on it enough to receive the survey.

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.

- a. What type of study are they conducting? Explain.

Experiment, they are imposing the treatment of the freezer or room temp.

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

Batteries were randomly selected and randomly assigned to treatments.

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

- a. What type of study should she conduct? Explain.

Either will work if justified. For example, Obs Study: Sample Survey to get their feedback about experience

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

Randomly select workers from several different businesses where people work in teams.

Secondary Math 3 Honors Note Guide 8.1

Drawing Conclusions from Data Unit

Basic Principles of Statistics

_____ is the _____ of _____.

- _____ are the objects (_____, _____, or _____) described by a set of data.
- A _____ is any _____ of an individual.
- A _____ variable places an individual into one of several _____ or _____.
- A _____ variable takes _____ values for which it _____ to find an _____.

Example: How can we help wood surfaces resist weathering, especially when restoring historic wooden buildings? In a study of this question, researchers prepared wooden panels and then exposed them to weather. Here are some of the variables recorded: type of wood (yellow poplar, pine, cedar); type of water repellent (solvent-based, water-based); paint thickness (millimeters); paint color (white, gray, light blue); weathering time (months). Identify each variable as categorical or quantitative.

type of wood

type of water repellent

paint thickness

paint color

weathering time

The _____ in a statistical study is the _____ group of individuals we want _____ about. A _____ collects data from _____ individual in the population. A _____ is a _____ of individuals in the population from which we _____ data.

- A _____ is a _____ that describes some _____ of the _____.
- A _____ is a _____ that describes some _____ of a _____.

For each bold number in the following descriptions, circle it if it is a statistic and underline it if it is a parameter.

Example: 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.

Example: 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 _____. Can we _____ the results? You will learn that it _____ on _____ the data were _____. It also becomes _____ to know _____ the data came from.

Types of Studies

An _____ observes individuals and _____ variables of _____ but does _____ attempt to _____ the responses.

A _____ uses an _____ to choose a sample that _____ some specific population, and then bases _____ about the population on _____ from the sample.

An _____ deliberately _____ some _____ on individuals to _____ their responses.

The method of the _____ of a sample is a _____ to _____ observational studies. It is _____ that a sample truly _____ the populations from which it is _____. For this reason, it is important to know how to _____ sampling _____ and how to sample _____.

How to Sample Poorly

The _____ of a statistical study shows _____ if it would consistently _____ or consistently _____ the _____ you want to know.

- _____ individuals from the _____ who are _____ results in a _____ sample.
 - A _____ sample consists of people who choose _____ by responding to a general _____.
- _____ of these sampling methods _____ from bias due to _____.

How to Sample Well

The best way to _____ problems of bias is to let _____ choose the sample. _____ involves using a _____ to determine which _____ of a population are _____ in the sample.

The most _____ method of random sampling is a _____. An SRS of size n is chosen in such a way that every _____ of n individuals in the population has an _____ to be selected as the sample. Examples include drawing names from a _____, using _____ or other _____, using a random number _____, or using random number _____.

Random Assignment in Experiments

A specific _____ applied to _____ in an experiment is called a _____. Ethically, a random sample _____ always be selected for _____ because _____ are imposed. Therefore to _____ bias in experiments, it is necessary to randomize in a different way than _____ participants. In an experiment, _____ means that _____ are assigned to _____ 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?
- 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.

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?
- 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.

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

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

- What type of study should she conduct? Explain.
- How might randomization be used in the design of the study you have selected?

