

Collecting Sample Data

Colby Community College

Definition

A **Placebo** is a treatment that has no medicinal effect. (Such as a sugar pill or saline injection.)

Definition

A **Placebo** is a treatment that has no medicinal effect. (Such as a sugar pill or saline injection.)

Example 1

In 1954, an experiment was designed to test the effectiveness of the Salk vaccine in preventing polio, which had killed or paralyzed thousands of children.

Definition

A **Placebo** is a treatment that has no medicinal effect. (Such as a sugar pill or saline injection.)

Example 1

In 1954, an experiment was designed to test the effectiveness of the Salk vaccine in preventing polio, which had killed or paralyzed thousands of children.

By random selection, 401,974 children were assigned to two groups:

- 200,745 children were given injections of the Salk vaccine.
- 201,229 children were given placebo injections that contained no drug.

Definition

A **Placebo** is a treatment that has no medicinal effect. (Such as a sugar pill or saline injection.)

Example 1

In 1954, an experiment was designed to test the effectiveness of the Salk vaccine in preventing polio, which had killed or paralyzed thousands of children.

By random selection, 401,974 children were assigned to two groups:

- 200,745 children were given injections of the Salk vaccine.
- 201,229 children were given placebo injections that contained no drug.

Among the children given the Salk vaccine, 33 later developed paralytic polio, and among the children given a placebo, 115 later developed paralytic polio.

Definition

In an **experiment**, we apply some treatment and then proceed to observe its effects on the individuals. (The individuals in experiments are called **subjects**.)

Definition

In an **experiment**, we apply some treatment and then proceed to observe its effects on the individuals. (The individuals in experiments are called **subjects**.)

Definition

In an **observational study**, we observe and measure specific characteristics, but we don't attempt to *modify* the individuals being studied.

Definition

In an **experiment**, we apply some treatment and then proceed to observe its effects on the individuals. (The individuals in experiments are called **subjects**.)

Definition

In an **observational study**, we observe and measure specific characteristics, but we don't attempt to *modify* the individuals being studied.

Note

In general, experiments are preferable to observational studies. But there are times where cost, time, or ethical concerns that prohibit the use of an experiment.

Definition

In an **experiment**, we apply some treatment and then proceed to observe its effects on the individuals. (The individuals in experiments are called **subjects**.)

Definition

In an **observational study**, we observe and measure specific characteristics, but we don't attempt to *modify* the individuals being studied.

Note

In general, experiments are preferable to observational studies. But there are times where cost, time, or ethical concerns that prohibit the use of an experiment.

Definition

A **lurking variable** is one that affects the variables included in the study, but is not included in the study.

Example 2

Suppose we want to determine if listening to music while driving increases the chance of being in an collision.

Example 2

Suppose we want to determine if listening to music while driving increases the chance of being in an collision.

- **Observational study:** If we gathered police reports about collisions and used them to determine if the person was listening to music or not.

Example 2

Suppose we want to determine if listening to music while driving increases the chance of being in an collision.

- **Observational study:** If we gathered police reports about collisions and used them to determine if the person was listening to music or not.
- **Experiment:** We randomly assign subjects to either listen to music while driving or listen to nothing. We then count how many collisions each subject is involved in.

Example 2

Suppose we want to determine if listening to music while driving increases the chance of being in an collision.

- **Observational study:** If we gathered police reports about collisions and used them to determine if the person was listening to music or not.
- **Experiment:** We randomly assign subjects to either listen to music while driving or listen to nothing. We then count how many collisions each subject is involved in.

Are there any lurking variables?

Example 2

Suppose we want to determine if listening to music while driving increases the chance of being in an collision.

- **Observational study:** If we gathered police reports about collisions and used them to determine if the person was listening to music or not.
- **Experiment:** We randomly assign subjects to either listen to music while driving or listen to nothing. We then count how many collisions each subject is involved in.

Are there any lurking variables?

Yes, automobile collisions happen due to a large number of reasons, many of which have nothing to do with music.

Experimental Design

- **Replication** is the repetition of an experiment on more than one individual. This means larger sample sizes are often needed.

Experimental Design

- **Replication** is the repetition of an experiment on more than one individual. This means larger sample sizes are often needed.
- **Blinding** is used when the subject doesn't know if they are receiving a placebo or a real treatment.

Experimental Design

- **Replication** is the repetition of an experiment on more than one individual. This means larger sample sizes are often needed.
- **Blinding** is used when the subject doesn't know if they are receiving a placebo or a real treatment.
 - **Double blind** is when both the patients and the researchers are unaware of who is getting the placebo and who is getting the treatment.

Experimental Design

- **Replication** is the repetition of an experiment on more than one individual. This means larger sample sizes are often needed.
- **Blinding** is used when the subject doesn't know if they are receiving a placebo or a real treatment.
 - **Double blind** is when both the patients and the researchers are unaware of who is getting the placebo and who is getting the treatment.
- **Randomization** is used when individuals are assigned to different groups through a process of random selection.

Experimental Design

- **Replication** is the repetition of an experiment on more than one individual. This means larger sample sizes are often needed.
- **Blinding** is used when the subject doesn't know if they are receiving a placebo or a real treatment.
 - **Double blind** is when both the patients and the researchers are unaware of who is getting the placebo and who is getting the treatment.
- **Randomization** is used when individuals are assigned to different groups through a process of random selection.

Definition

A **simple random sample** of n subjects is selected in such a way that every possible sample of the same size n has the same chance of being chosen.

Experimental Design

- **Replication** is the repetition of an experiment on more than one individual. This means larger sample sizes are often needed.
- **Blinding** is used when the subject doesn't know if they are receiving a placebo or a real treatment.
 - **Double blind** is when both the patients and the researchers are unaware of who is getting the placebo and who is getting the treatment.
- **Randomization** is used when individuals are assigned to different groups through a process of random selection.

Definition

A **simple random sample** of n subjects is selected in such a way that every possible sample of the same size n has the same chance of being chosen.

Note

While random sampling seems easy, it actually takes a great deal of planning to implement correctly.

Definition

Systematic sampling is where you select every k th subject.

Definition

Systematic sampling is where you select every k th subject.

Definition

Convenience sampling use data that is very easy to get.

Definition

Systematic sampling is where you select every k th subject.

Definition

Convenience sampling use data that is very easy to get.

Definition

Stratified sampling subdivides the population into groups with the same characteristics, then randomly sample within each group.

Definition

Systematic sampling is where you select every k th subject.

Definition

Convenience sampling use data that is very easy to get.

Definition

Stratified sampling subdivides the population into groups with the same characteristics, then randomly sample within each group.

Definition

Cluster sampling partitions the population into groups, then randomly selects some groups. All members of the chosen groups are included.

Definition

Systematic sampling is where you select every k th subject.

Definition

Convenience sampling use data that is very easy to get.

Definition

Stratified sampling subdivides the population into groups with the same characteristics, then randomly sample within each group.

Definition

Cluster sampling partitions the population into groups, then randomly selects some groups. All members of the chosen groups are included.

Definition

Multistage sampling is when multiple sampling methods are used to collect data.

Definition

Confounding occurs when we can see some effect, but we can't identify the specific factor that caused it.

Definition

Confounding occurs when we can see some effect, but we can't identify the specific factor that caused it.

Example 3

Let us assume that an experiment studying the effectiveness of a blood pressure drug grouped subjects by gender. The male subjects received the placebo treatment and the female subjects received the treatment. The researchers then recorded the blood pressure of all subjects.

Definition

Confounding occurs when we can see some effect, but we can't identify the specific factor that caused it.

Example 3

Let us assume that an experiment studying the effectiveness of a blood pressure drug grouped subjects by gender. The male subjects received the placebo treatment and the female subjects received the treatment. The researchers then recorded the blood pressure of all subjects.

If the data showed that the placebo group had, on average, higher blood pressure, does that mean the drug works?

Definition

Confounding occurs when we can see some effect, but we can't identify the specific factor that caused it.

Example 3

Let us assume that an experiment studying the effectiveness of a blood pressure drug grouped subjects by gender. The male subjects received the placebo treatment and the female subjects received the treatment. The researchers then recorded the blood pressure of all subjects.

If the data showed that the placebo group had, on average, higher blood pressure, does that mean the drug works?

We can't tell. It could be that the drug works, or it could be that men naturally have higher blood pressure than women. From the data collected, we just can't tell.

Definition

A **sampling error** occurs when the sample has been selected with a random method, but there is a discrepancy between a sample result and the true population result.

Definition

A **sampling error** occurs when the sample has been selected with a random method, but there is a discrepancy between a sample result and the true population result.

Definition

A **nonsampling error** is the result of human error, including such factors as wrong data entries, computing errors, biased questions, or applying inappropriate statistical methods.

Definition

A **sampling error** occurs when the sample has been selected with a random method, but there is a discrepancy between a sample result and the true population result.

Definition

A **nonsampling error** is the result of human error, including such factors as wrong data entries, computing errors, biased questions, or applying inappropriate statistical methods.

Definition

A **nonrandom sampling error** is the result of using a non-random sampling method, such as convenience sampling or voluntary response.

Definition

A **sampling error** occurs when the sample has been selected with a random method, but there is a discrepancy between a sample result and the true population result.

Definition

A **nonsampling error** is the result of human error, including such factors as wrong data entries, computing errors, biased questions, or applying inappropriate statistical methods.

Definition

A **nonrandom sampling error** is the result of using a non-random sampling method, such as convenience sampling or voluntary response.

Note

Experimental design is very important and very difficult. A full course in the design of experiments is needed to explore this topic fully.