Lesson 12: Types of Statistical Studies

Classwork

Opening Exercise

You want to know what proportion of the population likes rock music. You carefully consider three ways to conduct a study. What are the similarities and differences between the following three alternatives? Do any display clear advantages or disadvantages over the others?

a. You could pick a random sample of people and ask them the question, "Do you like rock music?" and record their answers.

b. You could pick a random sample of people and follow them for a period of time, noting their music purchases, both in stores and online.

c. You could pick a random sample of people, separate it into groups, and have each group listen to a different genre of music. You would collect data on the people who display an emotional response to the rock music.

A statistical study begins by asking a question that can be answered with data. The next steps are to collect appropriate data, organize and analyze it, and arrive at a conclusion in the context of the original question. This lesson focuses on the three main types of statistical studies: observational studies, surveys, and experiments. The objective of an observational study and a survey is to learn about characteristics of some population, so the data should be collected in a way that would result in a representative sample. This speaks to the importance of random selection of subjects for the study. The objective of an experiment is to answer such questions as, "What is the effect of treatments on a response variable?" Data in an experiment need to be collected in a way that does not favor one treatment over another. This demonstrates the importance of random assignment of subjects in the study to the treatments.

An observational study is one in which the values of one or more variables are observed with no attempt to affect the outcomes. One kind of observational study is a survey. A survey requires asking a group of people to respond to one or more questions. (A poll is one example of a survey.) An experiment differs from an observational study: In an experiment, subjects are assigned to treatments for the purpose of seeing what effect the treatment has on some response while an observational study makes no attempt to affect the outcomes, i.e., no treatment is given. Note that subjects could be people, animals, or any set of items that produce variability in their responses. Here is an example of an observational study: In a random sample of students, it was observed that those students who played a musical instrument had better grades than those who did not play a musical instrument. In an experiment, a group of students who do not currently play a musical instrument would be assigned at random to having to play a musical instrument or not having to play a musical instrument for a certain period of time. Then, at the end of the period of time we would compare academic performance.

Classify each of the three study methods about rock music as an observational study, a survey or an experiment.

Example 1: Survey

Item	I like the item.	I don't like the item.	I have never tried the item.
Salad			
Veg pizza			
Turkey sandwich			
Raspberry tea			

a. It is easy to determine if a study is a survey. A survey asks people to respond to questions. But surveys can be flawed in several ways. Questions may be confusing. For example, consider the following question:

What kind of computer do you own? (*Circle one*) Mac IBM-PC

How do you answer that question if you don't own a computer? How do you answer that question if you own a different brand? A better question would be:

Do you own a computer? (*Circle one*) Yes No

If you answered yes, what brand of computer is it? _____"

Now consider the question, "Do you like your school's cafeteria food?"

Re-write the question in a better form. Keep in mind that not all students may use the school's cafeteria, and even if they do, there may be some foods that they like and some that they don't like.

b. Something else to consider with surveys is how survey participants are chosen. If the purpose of the survey is to learn about some population, ideally participants would be randomly selected from the population of interest. If people are not randomly selected, misleading conclusions from the survey data may be drawn. There are many famous examples of this. Perhaps the most famous case was in 1936 when *The Literary Digest* magazine predicted that Alf Landon would beat incumbent President Franklin Delano Roosevelt by 370 electoral votes to 161. Roosevelt won 523 to 8.

Ten million questionnaires were sent to prospective voters (selected from the magazine's subscription list, automobile registration lists, phone lists, and club membership lists), and over two million questionnaires were returned. Surely such a large sample should represent the whole population. How could *The Literary Digest* prediction be so far off the mark?

c. Write or say to your neighbor two things that are important about surveys.

Example 2: Observational Study

a. An observational study records the values of variables for members of a sample but does not attempt to influence the responses. For example, researchers investigated the link between use of cell phones and brain cancer. There are two variables in this study: One is the extent of cell-phone usage, and the second is whether a person has brain cancer. Both variables were measured for a group of people. This is an observational study. There was no attempt to influence peoples' cell-phone usage to see if different levels of usage made any difference in whether or not a person developed brain cancer.

   Why would studying any relationship between asbestos exposure and lung cancer be an observational study and not an experiment?

b. In an observational study (just as in surveys), the people or objects to be observed would ideally be selected at random from the population of interest. This would eliminate bias and make it possible to generalize from a sample to a population. For example, to determine if the potato chips made in a factory contain the desired amount of salt, a sample of chips would be selected randomly so that the sample can be considered to be representative of the population of chips.

   Discuss how a random sample of 100 chips might be selected from a conveyor belt of chips.

c. Suppose that an observational study establishes a link between asbestos exposure and lung cancer. Based on that finding, can we conclude that asbestos exposure causes lung cancer? Why or why not?

d. Write or say to your neighbor two things that are important about observational studies.

Example 3: Experiment

a. An experiment imposes treatments to see the effect of the treatments on some response. Suppose that an observational study indicated that a certain type of tree did not have as much termite damage as other trees. Researchers wondered if resin from the tree was toxic to termites. They decided to do an experiment where they exposed some termites to the resin and others to plain water and recorded whether the termites survived. The explanatory variable (treatment variable) is the exposure type (resin, plain water), and the response variable is whether or not the termite survived. We know this is an experiment because the researchers imposed a treatment (exposure type) on the subjects (termites).

Is the following an observational study or an experiment? Why? If it is an experiment, identify the treatment variable and the response variable. If it is an observational study, identify the population of interest.

> A study was done to answer the question, "What is the effect of different durations of light and dark on the growth of radish seedlings?" Three similar growth chambers (plastic bags) were created in which $\mathbf{30}$ seeds randomly chosen from a package were placed in each chamber. One chamber was randomly selected and placed in $\mathbf{24}$ hours of light, another for $\mathbf{12}$ hours of light and $\mathbf{12}$ hours of darkness, and a third for $\mathbf{24}$ hours of darkness. After three days, researchers measured and recorded the lengths of radish seedlings for the germinating seeds.

b. In an experiment, random assignment of subjects to treatments is done to create comparable treatment groups. For example, a university biologist wants to compare the effects of two weed killers on pansies. She chooses 24 plants. If she applies weed killer A to the 12 healthiest plants and B to the remaining 12 plants, she won't know which plants died due to the type of weed killer used and which plants subjected to weed killer B were already on their "last legs." Randomly selecting twelve plants to receive weed killer A and then assigning the rest to B would help ensure that the plants in each group are fairly similar.

How might the biologist go about randomly assigning 12 plants from the 24 candidates to receive weed killer A? Could she be sure to get exactly 12 plants assigned to weed killer A and 12 plants to weed killer B by tossing a fair coin for each plant and assigning "heads up" plants to weed killer A and "tails up" to weed killer B? If not, suggest a method that you would use.

c. Write or say to your neighbor two things that are important about experiments.

Exercises 1–3

1. For each of the following study descriptions, identify whether the study is a survey, observational study, or experiment, and give a reason for your answer. For observational studies, identify the population of interest. For experiments, identify the treatment and response variables.

   a. A study investigated whether boys are quicker at learning video games than girls. Twenty randomly selected boys and twenty randomly selected girls played a video game that they had never played before. The time it took them to reach a certain level of expertise was recorded.

   b. As your statistics project, you collect data by posting five questions on poster board around your classroom and recording how your classmates respond to them.

   c. A professional sports team traded its best player. The local television station wanted to find out what the fans thought of the trade. At the beginning of the evening news program, they asked viewers to call one number if they favored the trade and a different number if they were opposed to the trade. At the end of the news program they announced that 53.7% of callers favored the trade.

   d. The local Department of Transportation is responsible for maintaining lane and edge lines on its paved roads. There are two new paint products on the market. Twenty comparable stretches of road are identified. Paint A is randomly assigned to ten of the stretches of road and paint B to the other ten. The department finds that paint B lasts longer.

   e. The National Highway Traffic Safety Administration conducts annual studies on drivers' seatbelt use at a random selection of roadway sites in each state in the United States. To determine if seatbelt usage has increased, data are analyzed over two successive years.

   f. People should brush their teeth at least twice a day for at least two to three minutes with each brushing. For a statistics class project, you ask a random number of students at your school questions concerning their tooth-brushing activities.

g. A study determines whether taking aspirin regularly helps to prevent heart attacks. A large group of male physicians of comparable health were randomly assigned equally to taking an aspirin every second day or to taking a placebo. After several years, the proportion of the males who had suffered heart attacks in each group was compared.

2. For the following, is the stated conclusion reasonable? Why or why not?

   A study found a positive relationship between the happiness of elderly people and the number of pets they have. Therefore, having more pets causes elderly people to be happier.

3. A researcher wanted to find out whether higher levels of a certain drug given to experimental rats would decrease the time it took them to complete a given maze to find food.

   a. Why would the researcher have to carry out an experiment rather than an observational study?

   b. Describe an experiment that the researcher might carry out based on 30 comparable rats and three dosage levels: 0 mg, 1 mg, and 2 mg.

Lesson Summary

- There are three major types of statistical studies: observational studies, surveys, and experiments.
  - An *observational study* records the values of variables for members of a sample.
  - A *survey* is a type of observational study that gathers data by asking people a number of questions.
  - An *experiment* assigns subjects to treatments for the purpose of seeing what effect the treatments have on some response.
- To avoid bias in observational studies and surveys, it is important to select subjects randomly.
- Cause and effect conclusions cannot be made in observational studies or surveys.
- In an experiment, it is important to assign subjects to treatments randomly in order to make cause-and-effect conclusions.

Problem Set

1. State if the following is an observational study, survey, or experiment, and give a reason for your answer.

   Linda wanted to know if it is easier for students to memorize a list of common three-letter words (such as fly, pen, red, …) than a list of three-letter nonsense words (such as vir, zop, twq, …). She randomly selected 28 students from all tenth-graders in her district. She put 14 blue and 14 red chips in a jar, and without looking each student chose a chip. Those with red chips were given the list of common words; those with blue chips were given the list of nonsense words. She gave all students one minute to memorize their list. After the minute, she collected the lists and asked them to write down all the words that they could remember. She recorded the number of correct words recalled.

2. State if the following is an observational study, survey, or experiment, and give a reason for your answer.

   Ken wants to compare how many hours a week that sixth graders spend doing mathematics homework to how many hours a week that eleventh graders spend doing mathematics homework. He randomly selects ten sixth graders and ten eleventh graders and records how many hours each student spent on mathematics homework in a certain week.

3. Suppose that in your health class you read two studies on the relationship between eating breakfast and success in school for elementary school children. Both studies concluded that eating breakfast causes elementary school children to be successful in school.
   a. Suppose that one of the studies was an observational study. Describe how you would recognize that they had conducted an observational study. Were the researchers correct in their causal conclusion?
   b. Suppose that one of the studies was an experiment. Describe how you would recognize that they had conducted an experiment. Were the researchers correct in their causal conclusion?

4. Data from a random sample of 50 students in a school district showed a positive relationship between reading score on a standardized reading exam and shoe size. Can it be concluded that having bigger feet causes one to have a higher reading score? Explain your answer.

Use the following scenarios for Problems 5–7.

A. Researchers want to determine if there is a relationship between whether or not a woman smoked during pregnancy and the birth weight of her baby. Researchers examined records for the past five years at a large hospital.

B. A large high school wants to know the proportion of students who currently use illegal drugs. Uniformed police officers asked a random sample of 200 students about their drug use.

C. A company develops a new dog food. The company wants to know if dogs would prefer its new food over the competition's dog food. One hundred dogs, who were food-deprived overnight, were given equal amounts of the two dog foods; the new food vs. competitor's food. The proportion of dogs preferring the new food was recorded.

5. Which scenario above describes an experiment? Explain why.

6. Which scenario describes a survey? Will the results of the survey be accurate? Why or why not?

7. The remaining scenario is an observational study. Is it possible to perform an experiment to determine if a relationship exists? Why or why not?