Carnegie Foundation for the Advancement of Teaching

1.1.3

Research Questions and Types of Statistical Studies

INTRODUCTION

In Lesson 1.1.1, we studied the four step process used in a statistical analysis. Step 1 in this process is "Ask a question that can be answered by collecting data." You will see in this lesson that understanding the type of research question being asked is very important. The type of research question has an impact on the method we use to collect data.

In earlier lessons you learned that a **population**, in a **statistical study**, is a set of all people or objects that share certain characteristics. A **sample** is a subset of the population used in a statistical study. **Subjects** are the individuals or objects in the study. The subjects are often people, but can be animals, plants, or things. **Variables** are the characteristics of subjects that we study. For example, a variable might be eye color, age, educational level, salary or city. In Lesson 1.1.1, we asked whether a relationship existed between a person's personality traits and birth date. The population was all people. The sample was the students in our class and the variables were sets of personality traits and birth date groups.

The questions we ask in a statistical study are usually in one of two categories:

1. Research questions about a population.
2. Research questions about the **cause-and-effect relationship** between two variables. Note that a cause-and-effect relationship is a relationship between two *variables* where a known change in the first variable produces a predictable change in the second.

Research Questions about a Population

Type of Research Question	Examples
Make an estimate about the population (often about an average or proportion)	• What is the average amount of sleep community college students get at night? • What percent of community college students have jobs?
Test a claim about the population (often a claim about an average or proportion)	• Is the average amount of sleep for community college students more than seven hours? • Do more than half of community college students plan to vote in November?
Compare two populations (often comparing averages or proportions)	• Do freshman university students have higher average GPA's than freshman community college students? • Are community college students more likely to receive financial aid than university students?
Investigate a relationship between two variables	• Is there a relationship between the number of hours a full time student works at a job and his or her GPA? • Are students who drink diet soda more likely to be overweight?

Research Questions about Cause-and-Effect Relationships between Variables

Examples of this type of research question include:

- Does requiring students to do homework in a college class improve test grades?
- Does caffeine reduce the risk of dementia (memory loss associated with old age)?
- Does taking an aspirin daily reduce the risk of heart attacks in adults over the age of 50?

To answer these research questions, we investigate how one variable responds as another variable is manipulated or changed. An **explanatory variable** is the variable being modified or manipulated in the study. A **response variable** is the output variable which is used to measure the impact of changes to the explanatory variable. An **experiment** involves a change to the explanatory variable.

TRY THESE

1 Here is one of the studies we examined in Lesson 1.1.1. As you read the problem, think about the following statistical ideas: (1) the population of interest, (2) the variables being studied, and (3) the type of research question that is being asked.

> A group of researchers studied women who had visited a fertility clinic. The researchers wondered if fewer than half of the women who visit the clinic would want to choose the sex of their future child.[1]
>
> The researchers mailed a survey to women who had visited the clinic. The survey asked women if they would choose the sex of their future child, if they were able to do so. A total of 561 women responded to the survey. Of these 561 women, 229 said that they wanted to choose the sex of their future child.
>
> The researchers did statistical analysis on the data. Based on their statistical analysis of these data, the researchers concluded that there is **convincing evidence** that *fewer than half* of the women who visit the clinic would choose the sex of a future child. This conclusion is based on the following observation:
>
> > If, in reality, at least half of women who visit a fertility clinic would like to choose the sex of a future child, it would be very unusual to observe a percentage as low as 41% in a sample of 561 women who visited the clinic (229/561 ≈ 0.41).

A What is the research question being asked?

B Does this study ask a question about a population or about a cause-and-effect relationship between two variables?

C What is the population of interest in this study?

D What variable is being examined for each subject in the study?

[1]Tarun Jain et al., "Preimplantation Sex Selection Demand and Preferences in an Infertility Population," *Fertility and Sterility* 83, no. 3 (2005): 649-58.

2 Here is the other study we examined in Lesson 1.1.1. As you read the problem, think about the following statistical ideas: (1) the population of interest, (2) the variables being studied, and (3) the type of research question that is being asked.

Researchers wanted to know if people think a task will be hard to accomplish when the instructions are difficult to read.[2] To answer this question, researchers randomly divided twenty student volunteers into two groups of 10 students each. Researchers gave instructions to each group of students using different fonts (see below). Instructions for one group were written in a large upright font. The other group was given the *same* instructions but in a font that used *hard-to-read italics*. Researchers asked students to read the directions and say how many minutes they thought the task would take. Researchers did this in order to figure out if the fonts used for the instructions made a difference.

This is the easy-to-read upright font that was used in the study.

This is the hard-to-read italic font that was used in the study.

The first group of students, those that read the instructions printed in the easy font, had an average time estimate of 8.23 minutes. The other group, the group that read the instructions in the *hard-to-read italic* font, had an average time estimate of 15.1 minutes.
Researchers concluded that such a large difference between the averages was not likely to have occurred by chance. There was evidence that people think a task will be harder when instructions are difficult to read.

A What is the research question being asked?

B Does this study ask a question about a population or about a cause-and-effect relationship between two variables?

C What is the explanatory variable?

D What is the response variable?

[2]Hyunjin Song, "The Effects of Processing Fluency on Judgment and Processing Style: Three Essays on Effort Prediction, Risk Perception, and Distortion Detection" (PhD diss., The University of Michigan, 2009).

NEXT STEPS

When we know what type of question a study asks and what the variable(s) are in the study, then we can move on to the second step of a statistical analysis. Step 2 in this four-step process is "decide what to measure and then collect data." There are two main types of studies used to collect data: **observational studies** and **experiments**.

Observational Study

In an *observational study*, researchers observe subjects in a sample to learn about population characteristics. Researchers usually observe a sample of the population, since it is often impossible to obtain information from every member of the population.

Because the goal of an observational study is typically to learn about the population, it is important that the sample be **representative** of the population.

Language Tip

A sample is *representative* of the population if it doesn't differ in any important way from the population.

The first study, about women choosing the sex of their child, is an observational study. The researchers only asked the women about choosing the sex. They were *observing* the women's responses.

Experiment

An *experiment* is used to answer questions about how one variable responds when another is manipulated. In an experiment, researchers observe how a response variable behaves as an explanatory variable is changed. Researchers actively manipulate the explanatory variable.

Because the goal of an experiment is to learn about the effect of the different treatments (changes to explanatory variable) on the response variable, it is important to apply the different treatments to similar samples. Also, the results only apply to a larger population if the sample is representative of the population.

The second study above is an experiment. The researchers changed the font of the instructions for a task to see if it would change the amount of time the students thought the task would take.

Key Difference between an Observational Study and an Experiment

An important difference between these data collection methods is that there is no attempt to influence the results in an observational study. This is different from an experiment. In an experiment, the explanatory variable is manipulated and changes in the response variable are examined.

TRY THESE

Read the two statistical studies below and answer the following questions.

3 Imagine that our college is having financial problems. The college announces that it will shorten library hours to save money. The library will be closed nights and weekends. Some students think that it is okay to pay a $20 fee to the college to keep the library open nights and weekends.

We are interested in learning about the proportion of students who would pay a $20 fee to keep the library open nights and weekends. To investigate this question, we select a sample of 100 students. We ask each of the students whether he or she agrees with the $20 fee increase to keep the library open nights and weekends.

A Does this study ask a question about a population or about a cause-and-effect relationship between two variables?

B Is this an observational study or an experiment?

If it is an *observational study*, what is the population of interest? What is the question we are asking about the population?

If it is an *experiment*, what is the explanatory variable and what is the response variable?

C Suppose that we collect data by asking 100 students who are entering the library whether they would pay the fee. Would this sample be representative of all students on campus? Why or why not?

D Suppose that we collect data by asking 100 students who we find hanging out in the school gym if they would pay the fee. Would this sample be representative of all students on campus? Why or why not?

E It is important to obtain a sample of 100 students that is representative of students at the college. In questions 3C and 3D, you read about some ways to select 100 students. Now think of a better way to select 100 students than the ideas in questions 3C and 3D. Why do you think your way is better? When you answer the question, be sure to think about (1) the population of interest, (2) your sample of 100 students, and (3) the type of research question that is being asked.

4 We are interested in learning whether jogging for longer amounts of time decreases the resting heart rate of college students. We want to see if there is a difference between:

- The resting heart rate of college students that jog for 30 minutes three times a week for six weeks, and
- The resting heart rate of college students that jog for 15 minutes three times a week for six weeks.

To investigate this question, we will use 100 college students who do not currently jog and who have volunteered to participate as subjects in this study. Resting heart rate of each subject will be measured at the start of the study. Fifty of the students will participate in a jogging program where they get together three times a week and jog for 30 minutes. The other 50 students will get together three times a week, but will only jog for 15 minutes. At the end of six weeks, resting heart rate will be measured again.

A What is the research question?

B Does this study ask a question about a population or about a cause-and-effect relationship between two variables?

C Is this an observational study or an experiment?

If it is an *observational study*, what is the population of interest? What is the question we are asking about the population?

If it is an *experiment*, what is the explanatory variable and what is the response variable?

D Imagine that we create the two groups for this study according to age. We group the 50 youngest volunteers in the 30 minute jogging group. We group the 50 oldest volunteers in the 15 minute jogging group. Is this a good idea? Why or why not?

E Imagine that we create the two groups for this study according to weight. We group the 50 volunteers that weigh the most into the 30 minute jogging group. We group the 50 volunteers that weigh the least into the 15 minute jogging group. Is this a good idea? Why or why not?

F We need to divide the 100 volunteers into two groups so that there is a "fair" comparison between the 30 minute and 15 minute jogging groups. What would be a better way to divide the 100 volunteers into two groups? Why would your way be better than to divide the volunteers by age (as in Question 4D) or weight (as in Question 4E)?

INTRODUCTION

Drawing Conclusions from Statistical Studies

The fourth step in the statistical process is **drawing a conclusion and communicating the results**. When making a conclusion, researchers extend beyond the data that are observed to explain what they learned from the study.

There are two types of conclusions that might be made from a study. One type of conclusion is "**generalizing from a sample to the population**." Note that researchers always seek to study a **representative sample** of a larger population. When researchers draw this type of conclusion, they are confident that what they observed in the sample is true for the larger population.

The other type of conclusion is "**cause-and-effect**." This conclusion arises from an experiment when a change in a response variable was caused by the manipulation of an explanatory variable. If a researcher manipulates a variable and this change generates an "effect" or response, the researcher can conclude that the change was due to the variable manipulation.

Language Tip

A conclusion is *reasonable* when the study design allows us to make that type of conclusion.

The table below summarizes when each of these types of conclusions is **reasonable**.

Type of Conclusion	Reasonable When
Generalize from sample to population	Observational study is conducted and the sample is representative of the population
Cause-and-effect	Experiment is conducted and groups assigned to experimental conditions are similar

The best way to choose a sample that is representative of the population is to use a **random sample** from the entire population. The best way to ensure similar groups for different experimental conditions is to use random assignment to the experimental groups. In addition, if an experiment uses both random assignment and a random sample from a population, we can draw a cause-and-effect conclusion and apply it to the larger population.

We will see more about these ideas in upcoming lessons, but without a random sample in an observational study or random assignment in an experiment, no conclusions can reliably be drawn.

In Summary:

- We can ask two types of research questions. Each type of research question is answered by a different type of study. **Observational studies** are used to answer research questions about characteristics of populations or comparing two populations. **Experiments** are used to answer research questions about cause and effect.
    - For observational studies, we should avoid making cause-and-effect conclusions, but it is possible to generalize from the sample to the population of interest if the study design incorporated random selection from the population.
    - For experiments, it is possible to reach cause-and-effect conclusions if the study design uses random assignment to create the experimental groups.
    - If an experiment uses both random assignment to create experimental groups and random selection from some population, it is possible to make cause-and-effect conclusions and to generalize these conclusions to the population.

5 Think about the study of women who visit a fertility clinic. Look at the design of the study. Is it reasonable to conclude that fewer than half of all women who visit the fertility clinic would choose the sex of their child? Explain why this is or is not a reasonable conclusion.

6 Think about the study about font style and the amount of time the reader thinks the task will take. Is it reasonable to conclude that a font which is more difficult to read will cause a reader to think the task described in the instructions will take longer? Explain why this is or is not a reasonable conclusion.

TRY THESE

7 The SAT exam is used in admissions decisions by many four-year colleges and universities. In 2006, The College Board carried out a study of 6,498 SAT essays that were selected at random from the more than 1.4 million SAT exams taken in the 2005 - 2006 academic year.[3] For this sample of essays, 15% were written in cursive and 85% were printed in block letters. The results showed that the average score for essays written in cursive was higher than the average score for essays that were printed.

A Is this study an observational study or an experiment? Explain.

B Is it reasonable to conclude that writing the essay in cursive was the cause of the higher scores? Explain your answer.

8 Imagine that a psychologist was interested in finding out whether music affects the ability to remember material that has been read. The psychologist recruited volunteer students who said they liked to study while listening to music and randomly assigned them into two groups. Each group was told to read an essay about Pearl Harbor and the U.S. entry into World War II. One group read the essay in silence. The other group read the essay while music of a style of their choice played in the background. After reading the essay they took a brief test that asked the students to recall details about the essay. The psychologist concluded that students who listened to music while they were reading scored lower than students who read in silence.

A Is this study an observational study or an experiment? Explain.

B What are two possible reasons for why the students who listened to music scored lower than the students who didn't listen to music?

[3]Krista D. Mattern, Wayne Camara, and Jennifer L. Kobrin, "SAT Writing: An Overview of Research and Psychometrics to Date," *collegeboard.org*, January 2007, https://research.collegeboard.org/sites/default/files/publications/2012/7/researchnote-2007-32-sat-writing-research-psychometrics.pdf.

C The psychologist found that the difference was so large that it was unlikely due to chance variation alone. Is it reasonable to conclude that the music was the cause of the lower scores? Explain.

9 An article described a study in which researchers looked at a random sample of 500 publicly accessible web profiles posted by 18-year-olds on a social networking site.[4] The content of each profile was analyzed. Researchers drew many conclusions. One of the conclusions was that profiles that list a sport or hobby have fewer references to "risky behavior." Fewer references to "risky behaviors" means that there were fewer comments on the person's profile about sex, drugs, or violence.

A Is it reasonable to generalize the stated conclusion to all 18-year-olds with a publicly accessible web profile on the social networking site? What aspect of the study supports your answer?

B Not all social networking site users have a publicly accessible profile. Is it reasonable to generalize the conclusion "profiles that list a sport or hobby have fewer references to risky behavior" to all 18-year-old users on the site? Why or why not?

10 The April 20, 2009 issue of the magazine *Sports Illustrated* reported the win-loss record for the Oklahoma City Thunder (a professional basketball team) for the 2008–2009 season. This record was actually worse for home games that were sold out (3 wins and 15 losses) than for home games that were not sold out (12 wins and 11 losses).

A Based on these data, is it reasonable to conclude that a sell-out crowd is the cause of the team's poor performance at sold-out home games? Can you think of another explanation for why the win-loss record might be worse for sold-out games than for games that are not sold out?

B Did random selection or random assignment play any role in the collection of these data?

[4]Megan A. Moreno et al., "Display of Health Risk Behaviors on MySpace by Adolescents: Prevalence and Associations," *Archives of Pediatrics & Adolescent Medicine* 163, no. 1 (2009): 27-34. http://www.ncbi.nlm.nih.gov/pubmed/19124700.

STUDENT NAME _____DATE _____

TAKE IT HOME

"Sweet Potatoes Brighten Your Skin" is the headline of an article that appeared in the magazine *Woman's World* (November 1, 2010).The article concludes that eating sweet potatoes causes skin to be healthier because it reverses age spots, blocks harmful UV rays in sunlight, and protects against skin dryness. Consider the following four hypothetical study designs. For each study design, answer the questions that follow.

1 **Study Design 1**: Two hundred students were selected at random from those enrolled at a large college in California. Each student in the sample was asked whether he or she ate sweet potatoes more than once in a typical week. A skin specialist rated skin health for each student on a scale of 1 to 10. It was concluded that skin health was significantly better on average for the group that reported eating sweet potatoes more than once a week than it was for the group that did not.

A Is the study described an observational study or an experiment? Explain your answer.

B Did the study use random selection from some population?

C Did the study use random assignment to experimental groups? If so, explain what the study did to randomly assign students?

D Is the conclusion "eating sweet potatoes leads to healthier skin" reasonable given the study description? Explain your answer.

E Is it reasonable to generalize conclusions from this study to some larger population? If so, what population?

2 **Study Design 2**: One hundred people who live in Miami volunteered to participate in a statistical study. The volunteers were divided into two experimental groups based on sex, with females in group 1 and males in group 2. Those in group 1 were asked to eat six ounces of sweet potatoes daily for three months. Those in group 2 were asked not to eat any sweet potatoes for three months. At the end of the three months, a skin specialist rated skin health on a scale of 1 to 10 for each of the volunteers. It was concluded that skin health was significantly better on average for group 1 than for group 2.

A Is the study described an observational study or an experiment? Explain your answer.

B Did the study use random selection from some population? If so, what did the study researchers do to randomly select students?

C Did the study use random assignment to experimental groups? If so, explain what the study did to randomly assign students?

D Is the conclusion "eating sweet potatoes leads to healthier skin" reasonable given the study description? Explain your answer.

E Is it reasonable to generalize conclusions from this study to some larger population? If so, what population?

3 **Study Design 3**: One hundred people volunteered to participate in a statistical study. For each volunteer, a coin was tossed in order to place them into a group. If the coin landed head up, the volunteer was assigned to group 1. If the coin landed tail up, the volunteer was assigned to group 2. Those in group 1 were asked to eat six ounces of sweet potatoes daily for three months. Those in group 2 were asked not to eat any sweet potatoes for three months. At the end of the three months, a skin specialist rated skin health on a scale of 1 to 10 for each of the volunteers. It was concluded that skin health was significantly better on average for those in group 1 than for those in group 2.

A Is the study described an observational study or an experiment? Explain your answer.

B Did the study use random selection from some population? If so, what did the study researchers do to randomly select students?

C Did the study use random assignment to experimental groups? If so, explain what the study did to randomly assign students?

D Is the conclusion "eating sweet potatoes leads to healthier skin" reasonable given the study description? Explain your answer.

E Is it reasonable to generalize conclusions from this study to some larger population? If so, what population?

4 **Study Design 4**: One hundred students were selected at random from those enrolled at a large college. Each of the selected students was asked to participate in a study and all agreed. For each student, a coin was tossed in order to place them into one of two groups. If the coin landed head up, the student was assigned to group 1. If the coin landed tail up, the student was assigned to group 2. Students in group 1 were asked to eat six ounces of sweet potatoes daily for three months. Students in group 2 were asked to not eat any sweet potatoes for three months. At the end of the three months, a skin specialist rated the health of each student's skin on a scale of 1 to 10. The researchers concluded that skin health was significantly better for the students in group 1 than for the students in group 2.

A Is the study described an observational study or an experiment? Explain your answer.

B Did the study use random selection from some population?

C Did the study use random assignment to experimental groups? If so, explain what the study did to randomly assign students?

D Is the conclusion “eating sweet potatoes leads to healthier skin” reasonable given the study description? Explain your answer.

E Is it reasonable to generalize conclusions from this study to some larger population? If so, what population?