PEARSON NEW INTERNATIONAL EDITION

Statistical Reasoning for Everyday Life

Jeff Bennett William L. Briggs
Mario F. Triola

Fourth Edition

ALWAYS LEARNING®

PEARSON®

Pearson New International Edition

Statistical Reasoning for Everyday Life

Jeff Bennett William L. Briggs
Mario F. Triola

Fourth Edition

PEARSON®

Pearson Education Limited
Edinburgh Gate
Harlow
Essex CM20 2JE
England and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsoned.co.uk

ISBN 10: 1-292-04021-1
ISBN 13: 978-1-292-04021-9

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Printed in the United States of America

Table of Contents

This page intentionally left blank

Speaking of Statistics

Is your drinking water safe? How many people approve of the President's budget plan? Are we getting good value for our health care dollars? Questions like these can be addressed only through statistical studies. In this chapter, we will discuss basic principles of statistical research and lay a foundation for the more detailed study of statistics. Along the way, we will consider a variety of examples that show how well-designed statistical studies can provide guidance for social policy and personal decisions, as well as a few cases in which statistics can be misleading.

Statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write.

—H. G. Wells

LEARNING GOALS

1. **What Is/Are Statistics?**
   Understand the two meanings of the term *statistics* and the basic ideas behind any statistical study, including the relationships among the study's population, sample, sample statistics, and population parameters.
2. **Sampling**
   Understand the importance of choosing a representative sample and become familiar with several common methods of sampling.
3. **Types of Statistical Studies**
   Understand the differences between observational studies and experiments; recognize key issues in experiments, including the selection of treatment and control groups, the placebo effect, and blinding.
4. **Should You Believe a Statistical Study?**
   Be able to evaluate statistical studies that you find in the media, so that you can decide whether the results are meaningful.

FOCUS TOPICS

Focus on Psychology: Are You Driving "Drunk" on Your Cell Phone?

Focus on Public Health: Is Your Lifestyle Healthy?

1 WHAT IS/ARE STATISTICS?

The subject of statistics is often stereotyped as dry or technical, but it touches on almost everything in modern society. Statistics can tell us whether a new drug is effective in treating cancer, it can help agricultural inspectors ensure that our food is safe, and it is essential for conducting and interpreting opinion polls. Businesses use statistics in market research and advertising. We even use statistics in sports, often as a way of ranking teams and athletes. Indeed, you'll be hard-pressed to think of any topic that is not linked with statistics in some important way.

The primary goal of this text is to help you learn the core ideas behind statistical methods. These basic ideas are not difficult to understand, although mastery of the details and theory behind them can require years of study. One of the great things about statistics is that even the small amount of theory covered in this text will give you the power to understand the statistics you encounter in the news, in your classes or workplace, and in your everyday life.

A good place to start is with the term *statistics* itself, which can be either singular or plural and has different meanings in the two cases. When it is singular, *statistics* is the *science* that helps us understand how to collect, organize, and interpret numbers or other information about some topic; we refer to the numbers or other pieces of information as *data*. When it is plural, *statistics* are the actual data that describe some characteristic. For example, if there are 30 students in your class and they range in age from 17 to 64, the numbers "30 students," "17 years," and "64 years" are all statistics that describe your class in some way.

BY THE WAY

Although you'll sometimes see the word *data* used as a singular synonym for *information*, technically it is plural: One piece of information is called a *datum*, and two or more pieces are called *data*.

Two Definitions of Statistics

- Statistics is the *science* of collecting, organizing, and interpreting data.
- Statistics are the *data* (numbers or other pieces of information) that describe or summarize something.

How Statistics Works

According to news reports, 111.3 million Americans watched the New York Giants win Super Bowl XLVI, which explains why the networks can now ask advertisers to pay more than $3 million for a 30-second commercial. But you may wonder: Who counted all these million people?

The answer is *no one*. The claim that 111.3 million people watched the Super Bowl came from statistical studies conducted by a company called Nielsen Media Research. This company compiles its famous *Nielsen ratings* by monitoring the television viewing habits of people in only about 5,000 homes.

BY THE WAY

Statistics originated with the collection of census and tax data, which are affairs of state. That is why the word *state* is at the root of the word *statistics*.

If you are new to the study of statistics, Nielsen's conclusion may seem like a stretch. How can anyone draw a conclusion about millions of people by studying just a few thousand? However, statistical science shows that this conclusion can be quite accurate, as long as the statistical study is conducted properly. Let's take the Nielsen ratings of the Super Bowl as an example and ask a few key questions that will illustrate how statistics works in general.

What Is the Goal of the Research?

Nielsen's goal is to determine the total number of Americans who watched the Super Bowl. In the language of statistics, we say that Nielsen is interested in the **population** of all Americans. The number that Nielsen hopes to determine—the number of people who watched the Super Bowl—is a particular characteristic of the population. In statistics, characteristics of the population are called **population parameters.**

Although we usually think of a population as a group of people, a statistical population can be any kind of group—people, animals, or things. For example, in a study of automobile safety, the population might be *all cars on the road*. Similarly, the term *population parameter* can refer to any characteristic of a population. In the case of automobile safety, the population parameters might include the total number of cars on the road during a certain time period, the accident rate among cars on the road, or the range of weights of cars on the road.

Definitions

The **population** in a statistical study is the *complete* set of people or things being studied.

Population parameters are specific numbers describing characteristics of the population.

EXAMPLE 1 Populations and Population Parameters

For each of the following situations, describe the population being studied and identify some of the population parameters that would be of interest.

a. You work for Farmers Insurance and you've been asked to determine the average amount paid to accident victims in cars without side-impact air bags.

b. You've been hired by McDonald's to determine the weights of the potatoes delivered each week for French fries.

c. You are a business reporter covering Genentech Corporation and you are investigating whether its new treatment is effective against childhood leukemia.

SOLUTION

a. The population consists of people who have received insurance payments for accidents in cars that lacked side-impact air bags. The relevant population parameter is the average amount paid to these people. (Later, the term "average" will be replaced by the more correct term "mean.")

b. The population consists of all the potatoes delivered each week for French fries. Relevant population parameters include the average weight of the potatoes and the variation of the weights (for example, are most of them close to or far from the average?).

c. The population consists of all children with leukemia. Important population parameters are the percentage of children who recover *without* the new treatment and the percentage of children who recover with the new treatment. • • ●

What Actually Gets Studied?

If researchers at Nielsen were all-powerful, they might determine the number of people watching the Super Bowl by surveying every individual American. But no one can do that, so instead they try to estimate the number of Americans watching by studying a relatively small group of people. Nielsen attempts to learn about the population of all Americans by carefully monitoring the viewing habits of a much smaller **sample** of Americans. More specifically, Nielsen uses recording devices in about 5,000 homes, so the people who live in these homes are the sample of Americans that Nielsen studies.

The individual measurements that Nielsen collects from the people in the 5,000 homes constitute the **raw data.** Nielsen collects much raw data—for example, when and how long each TV in the household is on, what show it is tuned to, and who in the household is watching. Nielsen then consolidates these raw data into a set of numbers that characterize the sample, such as the percentage of viewers in the sample who watched each individual television show or the total number of people in the sample who watched the Super Bowl. These numbers are called **sample statistics.**

Definitions

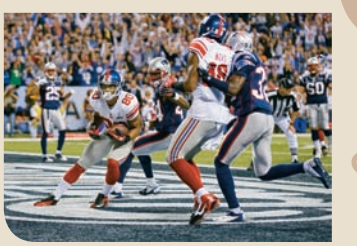
A **sample** is a subset of the population from which data are actually obtained.

The actual measurements or observations collected from the sample constitute the **raw data.**

Sample statistics are numbers describing characteristics of the sample found by consolidating or summarizing the raw data.

BY THE WAY

Arthur C. Nielsen founded his company and invented market research in 1923. He introduced the Nielsen Radio Index to rate radio programs in 1942 and extended his methods to television programming in the 1960s. The company now also tracks other media (Internet, smart phones, etc.) and must constantly adapt its methodology to new media technologies.



BY THE WAY

By the Labor Department definition, someone who is not working is not necessarily unemployed. For example, stay-at-home moms and dads are not counted among the unemployed unless they are actively trying to find a job, and people who tried to find work but gave up in frustration are not counted as unemployed.

EXAMPLE 2 Unemployment Survey

The U.S. Labor Department defines the *civilian labor force* as all those people who are either employed or actively seeking employment. Each month, the Labor Department reports the unemployment rate, which is the percentage of people actively seeking employment within the entire civilian labor force. To determine the unemployment rate, the Labor Department surveys 60,000 households. For the unemployment reports, describe each of the following.

a. population **b.** sample **c.** raw data **d.** sample statistics **e.** population parameters

SOLUTION

a. The *population* is the group that the Labor Department wants to learn about, which is all the people who make up the civilian labor force.

b. The *sample* consists of all the people among the 60,000 households surveyed.

c. The *raw data* consist of all the information collected in the survey.

d. The *sample statistics* summarize the raw data for the sample. In this case, the relevant sample statistic is the percentage of people in the sample who are actively seeking employment. (The Labor Department also calculates similar sample statistics for subgroups in the population, such as the percentages of teenagers, men, women, and veterans who are unemployed.)

e. The *population parameters* are the characteristics of the entire population that correspond to the sample statistics. In this case, the relevant population parameter is the actual unemployment rate. Note that the Labor Department does *not* actually measure this population parameter, because data are collected only for the sample and then are used to estimate the population parameter. • • ●

How Do Sample Statistics Relate to Population Parameters?

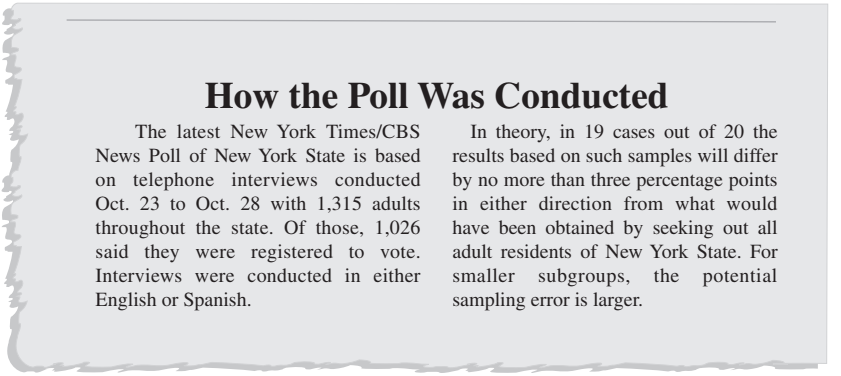
Suppose Nielsen finds that 31% of the people in the 5,000 homes in its sample watched the Super Bowl. This "31%" is a sample statistic, because it characterizes the sample. But what Nielsen really wants to know is the corresponding population parameter, which is the percentage of all Americans who watched the Super Bowl.

There is no way for Nielsen researchers to know the exact value of the population parameter, because they've studied only a sample. However, Nielsen researchers hope that they've done their work correctly so that the sample statistic is a good estimate of the population parameter. In other words, they would like to conclude that because 31% of the sample watched the Super Bowl, approximately 31% of the population also watched the Super Bowl. One of the primary purposes of statistics is to help researchers assess the validity of this type of conclusion.

TIME OUT TO THINK

Suppose Nielsen concludes that 30% of Americans watched the Super Bowl. How many people does this represent? (The population of the United States is approximately 310 million.)

Statistical science provides methods that enable researchers to determine how well a sample statistic estimates a population parameter. For example, results from surveys or opinion polls are usually quoted along with a value called the **margin of error.** By adding and subtracting the margin of error from the sample statistic, we find a range of values, or **confidence interval,** that is *likely* to contain the population parameter. In most cases, the margin of error is defined so that we can have 95% confidence that this range contains the population parameter; Figure 1 shows an explanation of this from the *New York Times*. In the case of the Nielsen ratings, the margin of error is about 1 percentage point. Therefore, if 31% of the sample was watching the Super Bowl, then we can be 95% confident that the range from 30% to 32% contains the actual percentage of the population watching the Super Bowl.



How the Poll Was Conducted

The latest New York Times/CBS News Poll of New York State is based on telephone interviews conducted Oct. 23 to Oct. 28 with 1,315 adults throughout the state. Of those, 1,026 said they were registered to vote. Interviews were conducted in either English or Spanish.

In theory, in 19 cases out of 20 the results based on such samples will differ by no more than three percentage points in either direction from what would have been obtained by seeking out all adult residents of New York State. For smaller subgroups, the potential sampling error is larger.

Figure 1 The margin of error in a survey or opinion poll usually describes a range that is likely (with 95% confidence, meaning in 19 out of 20 cases) to contain the population parameter. This excerpt from the *New York Times* explains a margin of error of 3 percentage points.

One of the most remarkable findings of statistical science is that it is possible to get meaningful results from surprisingly small samples. Nevertheless, larger sample sizes are better (when they are feasible), because the margin of error is generally smaller for larger samples. For example, the margin of error for a 95% confidence interval in a well-conducted poll is typically about 5 percentage points for a sample size of 400, but drops to 3 percentage points for a sample size of 1,000 and to 1 percentage point for a sample of 10,000.

Definition

The **margin of error** in a statistical study is used to describe the range of values, or **confidence interval,** likely to contain the population parameter. We find this confidence interval by adding and subtracting the margin of error from the sample statistic obtained in the study. That is, the range of values likely to contain the population parameter is

$$\begin{aligned} \text{from} \quad & (\text{sample statistic} - \text{margin of error}) \\ \text{to} \quad & (\text{sample statistic} + \text{margin of error}) \end{aligned}$$

The margin of error is usually defined to give a 95% confidence interval, meaning that 95% of samples of the size used in the study would contain the actual population parameter (and 5% would not).

EXAMPLE 3 Sex and Politics

The Pew Research Center for People and the Press interviewed 1,002 adult Americans and asked about the reason for a recent increase in sex scandals among elected officials. Fifty-seven percent of the respondents claimed that the increase is due to greater scrutiny by the media, while 19% felt that the increase is due to declining moral standards. The margin of error for the poll was 3 percentage points. Describe the population and the sample for this survey, and explain the meaning of the sample statistic of 57%. What can we conclude about the percentage of the population that believes the increase in political sex scandals is due to greater media scrutiny?

SOLUTION The population is all adult Americans and the sample consists of the 1,002 people who were interviewed. The sample statistic of 57% is the *actual* percentage of people in the sample who answered that greater media scrutiny is responsible for the increase in political sex scandals. The 57% sample statistic and the margin of error of 3 percentage points tell us that the range of values

$$\begin{aligned} \text{from} \quad & 57\% - 3\% = 54\% \\ \text{to} \quad & 57\% + 3\% = 60\% \end{aligned}$$

is likely (with 95% confidence) to contain the population parameter, which in this case is the true percentage of all adult Americans who believe that greater media scrutiny is responsible for the increase in political sex scandals. • • ●

TIME OUT TO THINK

In the poll described in Example 3, the respondents were given the two possible explanations, greater media scrutiny and lower moral standards. Do you think the results might have been different if respondents were asked to provide their own explanations? Explain.

Putting It All Together: The Process of a Statistical Study

The process used by Nielsen Media Research is similar to that used in many statistical studies. Figure 2 and the box below summarize the basic steps in a statistical study. Keep in mind that these steps are somewhat idealized, and the actual steps may differ from one study to another. Moreover, the details hidden in the basic steps are critically important. For example, a poorly chosen sample in Step 2 can render the entire study meaningless, and great care must be taken in inferring conclusions about a population from results found for the much smaller sample of that population.

BY THE WAY

Statisticians often divide their subject into two major branches: **descriptive statistics,** which deals with *describing* raw data in the form of graphics and sample statistics, and **inferential statistics,** which deals with *inferring* (or estimating) population parameters from sample data.

Basic Steps in a Statistical Study

Step 1. State the goal of your study precisely; that is, determine the population you want to study and exactly what you'd like to learn about it.

Step 2. Choose a representative sample from the population.

Step 3. Collect raw data from the sample, and summarize these data by finding sample statistics of interest.

Step 4. Use the sample statistics to make inferences about the population.

Step 5. Draw conclusions; determine what you learned and whether you achieved your goal.

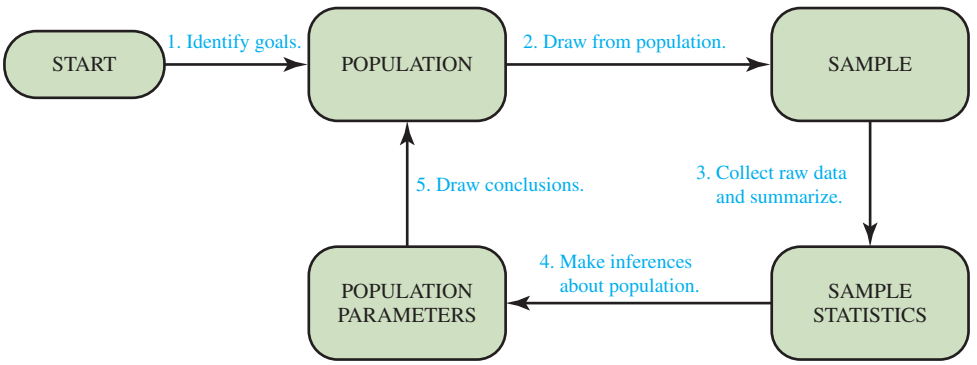


Figure 2 The process of a statistical study.

EXAMPLE 4 Identifying the Steps

Identify how researchers applied the five basic steps in the survey from Example 3.

SOLUTION The steps apply as follows.

1. The researchers had a goal of learning what Americans think about the causes of recent political scandals. They chose adult Americans as the population, deliberately leaving out children.
2. They chose 1,002 adult Americans for their sample. Although we are not told how the sample was drawn, we will assume that it was drawn so that the 1,002 adult Americans are typical of the entire adult American population.

3. They collected the raw data by asking a simple question of the people in the sample. The raw data are the individual responses to the question. They summarized these data with sample statistics, such as the overall percentages of people in the sample who chose each answer.
4. Techniques of statistical science allowed the researchers to infer population characteristics. In this case, the inference consisted of estimating the relevant population parameter and calculating the margin of error.
5. By making sure that the study was conducted properly and interpreting the estimates of the population parameters, the researchers drew overall conclusions about Americans' attitudes concerning recent scandals.

Statistics: Decisions for an Uncertain World

Most of the examples we've discussed so far involve surveys or polls, but the subject of statistics encompasses much more, including experiments designed to test new medical treatments, analyses of the dangers of global warming, and even assessments of the value of a college education. Indeed, it is fair to say that the primary purpose of statistics is to help us make good decisions whenever we are confronted with a variety of possible options.

The Purpose of Statistics

Statistics has many uses, but perhaps its most important purpose is to help us make good decisions about issues that involve uncertainty.

This purpose will be clear in most of the case studies and examples we consider in this text, but occasionally we'll have to discuss a bit of theory that may seem somewhat abstract at first. If you keep the overall purpose of statistics in mind, you'll be rewarded in the end when you see how the theory helps us understand our world. The following case study will give you a taste of what lies ahead. It involves several important theoretical ideas that led to one of the 20th century's greatest accomplishments in public health.

CASE STUDY The Salk Polio Vaccine

If you had been a parent in the 1940s or 1950s, one of your greatest fears would have been the disease known as polio. Each year during this long polio epidemic, thousands of young children were paralyzed by the disease. In 1954, a large experiment was conducted to test the effectiveness of a new vaccine created by Dr. Jonas Salk (1914–1995). The experiment involved a sample of 400,000 children chosen from the population of all children in the United States. Half of these 400,000 children received an injection of the Salk vaccine. The other half received an injection that contained only salt water. (The salt water injection was a *placebo*; see Section 3.) Among the children receiving the Salk vaccine, only 33 contracted polio. In contrast, there were 115 cases of polio among the children who did not get the Salk vaccine. Using techniques of statistical science that we'll study later, the researchers concluded that the vaccine was effective at preventing polio. They therefore decided to launch a major effort to improve the Salk vaccine and distribute it to the population of *all* children. Thanks to this vaccine (and improved ones developed later), the horror of polio is now largely a memory of the past.

BY THE WAY

Polio quickly became rare in the United States after the development of the Salk vaccine, but it remained common in less-developed countries. A global effort to vaccinate children against polio began in 1998 and has achieved great success, though it has not yet reached its goal of completely eradicating the disease.

The greatest reward for doing is the opportunity to do more.

—Jonas Salk

Section 1 Exercises

Statistical Literacy and Critical Thinking

1. **Population and Sample.** What is a *population*, what is a *sample*, and what is the difference between them?
2. **Statistic and Statistics.** Suppose that, in a discussion, one person refers to baseball *statistics* and another refers to the use of *statistics* in showing that a particular drug is an effective treatment. Do both uses of the term *statistics* have the same meaning? If not, how do they differ?
3. **Statistics and Parameters.** What is a sample statistic, what is a population parameter, and what is the difference between them?
4. **Margin of Error.** What is the margin of error in a statistical study, and why is it important?

Does It Make Sense? For Exercises 5–10, decide whether the statement makes sense (or is clearly true) or does not make sense (or is clearly false). Explain clearly; not all of these have definitive answers, so your explanation is more important than your chosen answer.

5. **Statistics and Parameters.** My professor conducted a statistical study in which he was unable to measure any sample statistics, but he succeeded in determining the population parameters with a very small margin of error.
6. **Poor Poll.** A poll conducted two weeks before the election found that Smith would get 70% of the vote, with a margin of error of 3%, but he ended up losing the election anyway.
7. **Poll Certainty.** There is no doubt that Johnson won the election, because an exit poll showed that she received 54% of the vote and the margin of error is only 3 percentage points.
8. **Beating Nielsen.** A new startup company intends to compete with Nielsen Media Research by providing data with a larger margin of error for the same price.
9. **Depression Sample.** The goal of my study is to learn about depression among people who have suffered through a family tragedy, so I plan to choose a sample from the population of patients in support groups for loss of a spouse.
10. **New Product.** Our market research department surveyed 1,000 consumers on their attitude toward our new product. Because the people in this sample were so enthusiastic in their desire to purchase the product, we have decided to roll out a nationwide advertising campaign.

Concepts and Applications

Population, Sample, Statistic, and Parameter. Exercises 11–14 each describe a statistical study. In each case, identify the sample, the population, the sample statistic, and the population parameter.

11. **Smoking Poll.** In a Gallup poll of 1,018 adults in the United States, it was found that 22% smoked cigarettes in the past week.
12. **Birth Weights.** For 186 randomly selected babies, the average (mean) of their birth weights is 3,103 grams (based on data from "Cognitive Outcomes of Preschool Children with Prenatal Cocaine Exposure," by Singer et al., *Journal of the American Medical Association,* Vol. 291, No. 20).
13. **Garlic and Cholesterol.** In a test of the effectiveness of garlic for lowering cholesterol, 47 adult subjects were treated with Garlicin, which is garlic in a processed tablet form. Cholesterol levels were measured before and after the treatment. The changes in their levels of LDL cholesterol (in mg/dL) have an average (mean) of 3.2 (based on data from "Effect of Raw Garlic vs Commercial Garlic Supplements on Plasma Lipid Concentrations in Adults With Moderate Hypercholesterolemia," by Gardner et al, *Archives of Internal Medicine*, Vol. 167).
14. **Job Interview Mistakes.** In an Accountemps survey of 150 senior executives, 47% said that the most common job interview mistake is to have little or no knowledge of the company.

Identifying the Range of Values. In Exercises 15–18, use the given statistics and margin of error to identify the range of values (confidence interval) likely to contain the true value of the population parameter.

15. **Wake Up.** A Braun Research poll asked 1,000 office workers how they wake up in time for work; 60% of them said that they use an alarm clock. The margin of error of was 3 percentage points.
16. **Wash Up.** *USA Today* reported that among 6,028 adults observed in restrooms, 85% washed their hands. The margin of error was 1 percentage point.
17. **Claim to Wash Up.** In a Harris Interactive survey of 1,006 adults, 96% say that they wash their hands when in a public restroom. The margin of error was 3 percentage points.
18. **Body Temperatures.** One hundred and six adults are randomly selected and tested for their body temperatures. Based on that sample, researchers estimated that the average (mean) body temperature is 98.2^{o} F with a margin of error of 0.1^{o} F.
19. **Global Warming Poll.** A Pew Research Center poll asked 1,708 randomly selected adults whether "global warming is a problem that requires immediate government action." Results

showed that 55% of those surveyed said yes. The margin of error was 2 percentage points. Can a news reporter safely write that the majority (more than 50%) of people believe that immediate government action is required?

20. Nielsen Survey. In a survey of 25,047 Super Bowl viewers, 51% of them said that they enjoyed commercials more than the game. The margin of error is 0.6%. Can we conclude that the majority of Super Bowl viewers prefer commercials to the game? Why or why not?

21. Do People Lie About Voting? In a survey of 1,002 people, 701 (or 70%) said that they voted in the last presidential election (based on data from ICR Research Group). The margin of error was 3 percentage points. However, actual voting records show that only 61% of all eligible voters actually voted. Does this imply that people lied when they responded in the survey? Explain.

22. Why the Discrepancy? An Eagleton Institute poll asked men if they agreed with this statement: "Abortion is a private matter that should be left to women to decide without government intervention." Among the men who were interviewed by women, 77% agreed with the statement. Among the men who were interviewed by men, 70% agreed with the statement. Assuming that the discrepancy is significant, how might that discrepancy be explained?

Interpreting Real Studies. For each of Exercises 23–26, do the following:

a. Based on the given information, state what you think was the goal of the study. Identify a possible population and the population parameter of interest.

b. Briefly describe the sample, raw data, and sample statistic for the study.

c. Based on the sample statistic and the margin of error, identify the range of values (confidence interval) likely to contain the population parameter of interest.

23. Death Penalty. A Gallup poll asked 511 randomly selected adults if they favor the death penalty for a person convicted of murder, with 64% saying yes. The margin of error is 4 percentage points.

24. Prescription Drugs. A study of 3,005 adults ages 57 to 85 showed that 82% of them use at least one prescription drug. The margin of error is 2 percentage points (based on data from "Use of Prescription and Over-the-Counter Medications and Dietary Supplements Among Older Adults in the United States," by Qato, et al., *Journal of the American Medical Association*, Vol. 300, No. 24).

25. Super Bowl. In a recent Super Bowl, a Nielsen report found that 45% of the 9,000 surveyed U.S. households had TV sets tuned to the game. The margin of error is 1 percentage point.

26. Piercings and Tattoos. A Harris Interactive survey of 514 human resources professionals showed that 46% of them say that piercings or tattoos make job applicants less likely to be hired. The margin of error for the survey was 4 percentage points.

Five Steps in a Study. Describe how you would apply the five basic steps in a statistical study (as listed in the box shown previously) to the issues in Exercises 27–30.

27. Cell Phones and Driving. You want to determine the percentage of drivers who use cell phones while they are driving.

28. Credit Scores. FICO (Fair Isaac Corporation) scores are routinely used to rate the quality of consumer credit. You want to determine the average (mean) FICO score of an adult in the United States.

29. Passenger Weight. Recognizing that overloading commercial aircraft would lead to unsafe flights, you want to determine the average (mean) weight of airline passengers.

30. Pacemaker Batteries. Because the batteries used in heart pacemakers are so critically important, you want to determine the average (mean) length of time that such batteries last before failure.

PROJECTS FOR THE INTERNET & BEYOND

31. Current Nielsen Ratings. Find the Nielsen ratings for the past week. What were the three most popular television shows? Explain the meaning of the "rating" and the "share" for each show.

32. Nielsen Methods. Nielsen Media Research supplies statistical data that is very important to advertisers and marketers. Visit the Nielsen Web site and read about one or more of their data products. Write a brief report on how they collect the data and how the data are used.

33. Comparing Airlines. The U.S. Department of Transportation routinely publishes on-time performance, lost baggage rates, and other statistics for different airline companies. Find a recent example of such statistics. Based on what you find, is it fair to say that any particular airline stands out as better or worse than others? Explain.

34. Labor Statistics. Use the Bureau of Labor Statistics Web site to find monthly unemployment rates over the past 12 months. If you assume that the monthly survey has a margin of error of about 0.2 percentage point, has there been a noticeable change in the unemployment rate over the past year? Explain.

35. Statistics and Safety. Identify a study that has been done (or should be done) to improve the safety of car drivers and passengers. Briefly describe the importance of statistics to the study.

36. Pew Research Center. The Pew Research Center for the People and the Press studies public attitudes toward the press, politics, and public policy issues. Go to its Web site and find the latest survey about attitudes. Select a particular recent survey, and write a summary of what was surveyed, how the survey was conducted, and what was found.

IN THE NEWS

37. Statistics in the News. Identify three stories from the past week that involve statistics in some way. In each case, write a brief statement describing the role of statistics in the story.

38. Statistics in Your Major. Write a brief description of some ways in which you think that the science of statistics can be used in your major field of study. (If you have not yet selected a major, answer the question for a major that you are considering.)

39. Statistics and Entertainment. The Nielsen ratings are well known for their role in gauging television viewing. Identify another way that statistics are used in the entertainment industry. Briefly describe the role of statistics in this application.

40. Statistics in Sports. Choose a sport and describe at least three different statistics commonly tracked by participants in or spectators of the sport. In each case, briefly describe the importance of the statistic to the sport.

41. Economic Statistics. The government regularly publishes many different economic statistics, such as the unemployment rate, the inflation rate, and the surplus or deficit in the federal budget. Study recent newspapers and identify five important economic statistics. Briefly explain the purpose of each of these five statistics.

2 SAMPLING

Not everything that can be counted counts, and not everything that counts can be counted.

—Albert Einstein

The only way to know the true value of a population parameter is to observe *every* member of the population. For example, to learn the exact mean height of all students at your school, you'd need to measure the height of every student. A collection of data from every member of a population is called a **census.** Unfortunately, conducting a census is often impractical. In some cases, the population is so large that it would be too expensive or time-consuming to collect data from every member. In other cases, a census may be ruled out because it would interfere with a study's overall goals. For example, a study designed to test the quality of candy bars before shipping could not involve a census because that would mean testing a piece of every candy bar, leaving none intact to sell.

Definition

A **census** is the collection of data from *every* member of a population.

Fortunately, most statistical studies can be done without going to the trouble of conducting a census. Instead of collecting data from every member of the population, we collect data from a sample and use the sample statistics to make inferences about the population. Of course, the inferences will be reasonable only if the members of the sample represent the population fairly, at least in terms of the characteristics under study. That is, we seek a **representative sample** of the population.

Definition

A **representative sample** is a sample in which the relevant characteristics of the sample members are generally the same as the characteristics of the population.

EXAMPLE 1 A Representative Sample for Heights

Suppose you want to determine the mean height of all students at your school. Which is more likely to be a representative sample for this study: the men's basketball team or the students in your statistics class?

SOLUTION The men's basketball team is not a representative sample for a study of height, both because it consists only of men and because basketball players tend to be taller than average. The mean height of the students in your statistics class is much more likely to be close to the mean height of all students, so the members of your class make a more representative sample than the members of the men's basketball team. • • ●

Bias

Imagine that, for the 5,000 homes in its sample, Nielsen chose only homes in which the primary wage earners worked a late-night shift. Because late-night workers aren't home to watch late-night television, Nielsen would find late-night shows to be unpopular among the homes in this sample. Clearly, this sample would *not* be representative of all American homes, and it would be wrong to conclude that late-night shows were unpopular among all Americans. We say that such a sample is *biased* because the homes in the sample differed in a specific way from "typical" American homes. (In reality, Nielsen takes great care to avoid such obvious bias in the sample selection.) More generally, the term **bias** refers to any problem in the design or conduct of a statistical study that tends to favor certain results. We cannot trust the conclusions of a biased study.

Definition

A statistical study suffers from **bias** if its design or conduct tends to favor certain results.

Bias can arise in many ways. For example:

- A sample is biased if the members of the sample differ in some specific way from the members of the general population. In that case, the results of the study will reflect the unusual characteristics of the sample rather than the actual characteristics of the population.
- A researcher is biased if he or she has a personal stake in a particular outcome. In that case, the researcher might intentionally or unintentionally distort the true meaning of the data.
- The data set itself is biased if its values were collected intentionally or unintentionally in a way that makes the data unrepresentative of the population.
- Even if a study is done well, it may be reported in a biased fashion. For example, a graph representing the data may tell only part of the story or depict the data in a misleading way.

Preventing bias is one of the greatest challenges in statistical research. Looking for bias is therefore one of the most important steps in evaluating a statistical study or media reports about a statistical study.

BY THE WAY

Many medical studies are experiments designed to test whether a new drug is effective. In an article published in the *Journal of the American Medical Association*, the authors found that studies with positive results (the drug is effective) are more likely to be published than studies with negative results (the drug is not effective). This "publication bias" tends to make new drugs, as a group, seem more effective than they really are.

EXAMPLE 2 Why Use Nielsen?

Nielsen Media Research earns money by charging television stations and networks for its services. For example, NBC pays Nielsen to provide ratings for its television shows. Why doesn't NBC simply do its own ratings, instead of paying a company like Nielsen to do them?

SOLUTION The cost of advertising on a television show depends on the show's ratings. The higher the ratings, the more the network can charge for advertising—which means NBC would have a clear bias if it conducted its own ratings. Advertisers therefore would not trust ratings that NBC produced on its own. By hiring an independent source, such as Nielsen, NBC can provide information that advertisers are more likely to believe. • • ●

TIME OUT TO THINK

The fact that NBC pays Nielsen for its services might seem to give Nielsen a financial incentive to make NBC look good in the ratings. If you worked for an advertising agency, what other factors might help you gain confidence in Nielsen's ratings?

Sampling Methods

A good statistical study *must* have a representative sample. Otherwise the sample is biased and conclusions from the study are not trustworthy. Let's examine a few common sampling methods that, at least in principle, can provide a representative sample.

Simple Random Samples

In most cases, the best way to obtain a representative sample is by choosing *randomly* from the population. A **random sample** is one in which every member of the population has an equal chance of being selected to be part of the sample. For example, you could obtain a random sample by having everyone in a population roll a die and choosing those people who roll a 6. In contrast, the sample would not be random if you chose everyone taller than 6 feet, because not everyone would have an equal chance of being selected.

In statistics, we usually decide in advance the sample size that is needed. With **simple random sampling,** every possible sample of a particular size has an equal chance of being selected. For example, to choose a simple random sample of 100 students from all the students in your school, you could assign a number to each student in your school and choose the sample by drawing 100 of these numbers from a hat. As long as each student's number is in the hat only once, every sample of 100 students has an equal chance of being selected. As a faster alternative to using a hat, you might choose the student numbers with the aid of a computer or calculator that has a built-in *random number generator*.

TIME OUT TO THINK

Look for the random number key on a calculator. (Nearly all scientific calculators have one.) What happens when you push it? How could you use the random number key to select a sample of 100 students?

Because simple random sampling gives every sample of a particular size the same chance of being chosen, it is likely to provide a representative sample, as long as the sample size is large enough.

EXAMPLE 3 Local Resident Sampling

You want to conduct an opinion poll in which the population is all the residents in a town. Could you choose a simple random sample by randomly selecting names from local property tax records?

SOLUTION A sample drawn from property tax records is not a simple random sample of the town population because these records would only list people who own property in the town. These records are therefore missing many town residents, and might also include people who live elsewhere but own property in the town. • • ●

Systematic Sampling

Simple random sampling is effective, but in many cases we can get equally good results with a simpler technique. Suppose you are testing the quality of microchips produced by Intel. As the chips roll off the assembly line, you might decide to test every 50th chip. This ought to give a

representative sample because there's no reason to believe that every 50th chip has any special characteristics compared with other chips. This type of sampling, in which we use a system such as choosing every 50th member of a population, is called **systematic sampling.**

EXAMPLE 4 Museum Assessment

When the National Air and Space Museum wanted to test possible ideas for a new solar system exhibit, a staff member interviewed a sample of visitors selected by systematic sampling. She interviewed a visitor exactly every 15 minutes, choosing whoever happened to enter the current solar system exhibit at that time. Why do you think she chose systematic sampling rather than simple random sampling? Was systematic sampling likely to produce a representative sample in this case?

SOLUTION Simple random sampling might occasionally have selected two visitors so soon after each other that the staff member would not have had time to interview each of them. The systematic process of choosing a visitor every 15 minutes prevented this problem from arising. Because there's no reason to think that the people entering at a particular moment are any different from those who enter a few minutes earlier or later, this process is likely to give a representative sample of the population of visitors during the time of the sampling. • • ●

EXAMPLE 5 When Systematic Sampling Fails

You are conducting a survey of students in a co-ed dormitory in which males are assigned to odd-numbered rooms and females are assigned to even-numbered rooms. Can you obtain a representative sample when you choose every 10th room?

SOLUTION No. If you start with an odd-numbered room, every 10th room will also be odd-numbered (such as room numbers 3, 13, 23, …). Similarly, if you start with an even-numbered room, every 10th room will also be even-numbered. You will therefore obtain a sample consisting of either all males or all females, neither of which is representative of the co-ed population. • • ●

TIME OUT TO THINK

Suppose you chose every fifth room, rather than every 10th room, in Example 5. Would the sample then be representative?

Convenience Samples

Systematic sampling is easier than simple random sampling but might also be impractical in many cases. For example, suppose you want to know the proportion of left-handed students at your school. It would take great effort to select a simple random sample or a systematic sample, because both require drawing from all the students in the school. In contrast, it would be easy to use the students in your statistics class as your sample—you could just ask the left-handed students to raise their hands. This type of sample is called a **convenience sample** because it is chosen for convenience rather than by a more sophisticated procedure. For trying to find the proportion of left-handed people, the convenience sample of your statistics class is probably fine; there is no reason to think that there would be a different proportion of left-handed students in a statistics class than anywhere else. But if you were trying to determine the proportions of students with different majors, this sample would be biased because some majors require a statistics course and others do not. In general, convenience sampling tends to be more prone to bias than most other forms of sampling.

EXAMPLE 6 Salsa Taste Test

A supermarket wants to decide whether to carry a new brand of salsa, so it offers free tastes at a stand in the store and asks people what they think. What type of sampling is being used? Is the sample likely to be representative of the population of all shoppers?

BY THE WAY

The sampling described in Example 4 was undertaken prior to the construction of the Voyage Scale Model Solar System, a permanent scale model exhibit that stretches along the National Mall from the National Air and Space Museum to the Smithsonian Castle. The photo below shows the son of one of the exhibit creators (also an author of this text) touching the scale model Sun.

SOLUTION The sample of shoppers stopping for a taste of the salsa is a convenience sample because these people happen to be in the store and are willing to try the new product. (This type of convenience sample, in which people choose whether or not to be part of the sample, is also called a *self-selected sample*. We will study self-selected samples further in Section 4.) This sample is unlikely to be representative of the population of all shoppers, because different types of people may shop at different times (for example, stay-at-home parents are more likely to shop at midday than are working parents) and only people who like salsa are likely to participate. The data might be still be useful, however, because the opinions of people who like salsa are probably the most important ones in this case. • • ●

Cluster Samples

Cluster sampling involves the selection of *all* members in randomly selected groups, or *clusters*. Imagine that you work for the Department of Agriculture and wish to determine the percentage of farmers who use organic farming techniques. It would be difficult and costly to collect a simple random sample or a systematic sample because either would require visiting many individual farms that are located far from one another. A convenience sample of farmers in a single county would be biased because farming practices vary from region to region. You might therefore decide to select a few dozen counties at random from across the United States and survey *every* farmer in each of those counties. We say that each county contains a *cluster* of farmers, and the sample consists of *every* farmer within the randomly selected clusters.

EXAMPLE 7 Gasoline Prices

You want to know the mean price of gasoline at gas stations located within a mile of rental car locations at airports. Explain how you might use cluster sampling in this case.

SOLUTION You could randomly select a few airports around the country. For these airports, you would check the gasoline price at *every* gas station within a mile of the rental car location. • • ●

BY THE WAY

As mandated by the U.S. Constitution, voting for the President is actually done by a small group of people called *electors*. Each state may select as many electors as it has members of Congress (counting both senators and representatives). When you cast a ballot for President, you actually cast a vote for your state's electors, each of whom has promised to vote for a particular presidential candidate. The electors cast their votes a few weeks after the general election.

Stratified Samples

Suppose you are conducting a poll to predict the outcome of the next U.S. presidential election. The population under study is all likely voters, so you might choose a simple random sample from this population. However, because presidential elections are decided by electoral votes cast on a state-by-state basis, you'll get a better prediction if you determine voter preferences within each state. Your overall sample should therefore consist of separate random samples from each of the 50 states. In statistical terminology, the populations of the 50 states represent subgroups, or **strata,** of the total population. Because your overall sample consists of randomly selected members from each stratum, you've used **stratified sampling.**

EXAMPLE 8 Unemployment Data

The U.S. Labor Department surveys 60,000 households each month to compile its unemployment report (see Example 2 in Section 1). To select these households, the department first groups cities and counties into about 2,000 geographic areas. It then randomly selects households to survey within these geographic areas. How is this an example of stratified sampling? What are the strata? Why is stratified sampling important in this case?

SOLUTION The unemployment survey is an example of stratified sampling because it first breaks the population into subgroups. The subgroups, or strata, are the people in the 2,000 geographic regions. Stratified sampling is important in this case because unemployment rates are likely to differ in different geographic regions. For example, unemployment rates in rural Kansas may be very different from those in Silicon Valley. By using stratified sampling, the Labor Department ensures that its sample fairly represents all geographic regions. • • ●

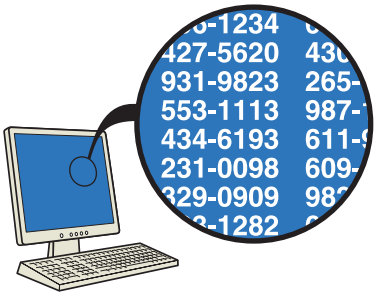
Summary of Sampling Methods

The following box and Figure 3 summarize the five sampling methods we have discussed. No single method is "best," as each one has its uses. (Some studies even combine two or more types of sampling.) But regardless of how a sample is chosen, keep in mind the following three key ideas:

- A study can be successful only if the sample is representative of the population.
- A biased sample is unlikely to be a representative sample.
- Even a well-chosen sample may still turn out to be unrepresentative just because of bad luck in the actual drawing of the sample.

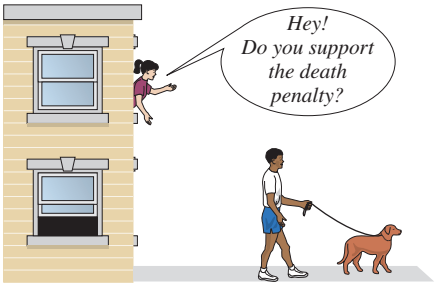
Common Sampling Methods

- **Simple random sampling:** We choose a sample of items in such a way that every sample of the same size has an equal chance of being selected.
- **Systematic sampling:** We use a simple system to choose the sample, such as selecting every 10th or every 50th member of the population.
- **Convenience sampling:** We use a sample that happens to be convenient to select.
- **Cluster sampling:** We first divide the population into groups, or clusters, and select some of these clusters at random. We then obtain the sample by choosing *all* the members within each of the selected clusters.
- **Stratified sampling:** We use this method when we are concerned about differences among subgroups, or *strata*, within a population. We first identify the strata and then draw a random sample within each stratum. The total sample consists of all the samples from the individual strata.



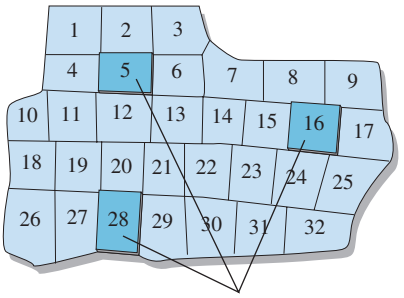
Simple Random Sampling:
Every sample of the same size has an equal chance of being selected. Computers are often used to generate random numbers.

Systematic Sampling:
Select every kth member.



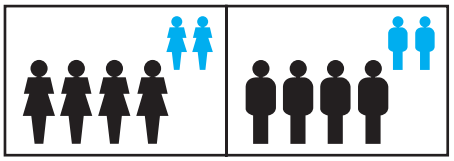
Convenience Sampling:
Use results that are readily available.

Election precincts in Carson County



Interview all voters in shaded precincts.

Cluster Sampling:
Divide the population into clusters, randomly select some of those clusters, then choose all members of the selected clusters.



Stratified Sampling:
Partition the population into at least two strata, then draw a sample from each.

Figure 3 Common sampling methods.

EXAMPLE 9 Sampling Methods

Identify the type of sampling used in each of the following cases.

a. The apple harvest from an orchard is collected in 1,200 baskets. An agricultural inspector randomly selects 25 baskets and then checks every apple in each of these baskets for worms.

b. An educational researcher wants to know whether, at a particular college, men or women tend to ask more questions in class. Of the 10,000 students at the college, she interviews 50 randomly selected men and 50 randomly selected women.

c. In trying to learn about planetary systems, astronomers conduct a survey by looking for planets among 100 nearby stars.

d. To determine who will win autographed footballs, a computer program randomly selects the ticket numbers of 11 people in a stadium filled with people.

SOLUTION

a. The apple inspection is an example of cluster sampling because the inspector begins with a randomly selected set of clusters (baskets) and then checks every apple in the selected clusters.

b. The groups of men and women represent two different strata for this study, so this is an example of stratified sampling.

c. The astronomers presumably focus on nearby stars because they are easier to study in this case, so this is an example of convenience sampling.

d. Because the computer selects the 11 ticket numbers at random, every ticket number has an equal chance of being chosen. This is an example of simple random sampling. • • ●

Section 2 Exercises

Statistical Literacy and Critical Thinking

1. **Census and Sample.** What is a *census,* what is a sample, and what is the difference between them?

2. **Biased Sample.** In a survey before a general election (as opposed to a primary), a sample is drawn randomly from a list of registered Democrats. Is there anything wrong with this sampling method?

3. **Cluster and Stratified Sampling.** Cluster sampling and stratified sampling both involve selecting subjects in subgroups of the population. What is the difference between those two types of sampling?

4. **Sample of Students.** A college statistics teacher conducted a study by recording whether each student in his class was right-handed. The objective was to form a conclusion about the proportion of college students that is right-handed. What type of sample was obtained? Is this sample likely to be biased? Why or why not?

Does It Make Sense? For Exercises 5–8, decide whether the statement makes sense (or is clearly true) or does not make sense (or is clearly false). Explain clearly; not all of these have definitive answers, so your explanation is more important than your chosen answer.

5. **Graduation Age.** For a statistics class project, I conducted a census to determine the mean age of students when they earn their bachelor's degrees.

6. **Convenience Sample.** For a statistics class project, I used a convenience sample, but the results may still be meaningful.

7. **Biased Sample.** The study must have been biased, because it concluded that 75% of Americans are more than 6 feet tall.

8. **Death Row.** There are currently 3,242 convicts on death row (based on 2010 data from the Bureau of Justice Statistics). We obtained a simple random sample of those convicts by compiling a numbered list, then using a computer to randomly generate 20 numbers between 1 and 3,242, then selecting the convicts that correspond to the generated numbers.

Concepts and Applications

Census. In Exercises 9–12, determine whether a census is practical in the situations described. Explain your reasoning.

9. **Laker Heights.** You want to determine the mean height of all basketball players on the LA Lakers team.

10. **High School Heights.** You want to determine the mean height of all high school basketball players in the United States.

11. **IQ Scores.** You want to determine the mean IQ score of all statistics instructors in the United States.

12. **Instructor Ages.** You want to determine the mean age of all statistics instructors at the University of Colorado.

Representative Samples? In Exercises 13–16, identify the sample, population, and sampling method. Then comment on whether you think it is likely that the sample is representative of the population.

13. **Senate Terms.** A political scientist randomly selects 4 of the 100 senators currently serving in Congress, then finds the lengths of time that they have served.

14. **Super Bowl.** During the Super Bowl game, Nielsen Media Research conducts a survey of 5,108 randomly selected households and finds that 44% of them have television sets tuned to the Super Bowl.

15. **Gun Ownership.** In a Gallup poll of 1,059 randomly selected adults, 39% answered "yes" when asked "Do you have a gun in your home?"

16. **Mail Survey.** A graduate student conducts a research project about how adult Americans send thank-you notes. She uses the U.S. Postal Service to mail a survey to 500 adults that she knows and asks them to mail back a response to this question: "Do you prefer to send thank you notes by e-mail or snail mail (the U.S. Postal Service)?" She gets back 65 responses, with 42 of them indicating a preference for snail mail.

Evaluate the Sample Choices. Exercises 17–18 each describe the goal of a study, then offer you four options for obtaining a sample. In each case, decide which sample is most likely to be a representative sample, and explain why. Then explain why each of the other choices is *not* likely to make a representative sample for the study.

17. **Credit Card Debt.** You want to determine the mean amount of credit card debt owed by adult consumers in Florida.

    - Sample 1: The Florida drivers who own and have registered Land Rover vehicles
    - Sample 2: The first 1,000 Florida residents listed in the Fort Lauderdale phone book
    - Sample 3: The first 1,000 Florida residents in a complete list of all Florida telephone numbers
    - Sample 4: The Florida residents who mail back a survey printed in the *Miami Herald*

18. **California Voters.** You want to conduct a survey to determine the proportion of eligible voters in California likely to vote for the Democratic presidential candidate in the next election.

    - Sample 1: All eligible voters in San Diego County
    - Sample 2: All eligible voters in the city of Sonoma
    - Sample 3: All eligible voters who respond to a CNN Internet survey
    - Sample 4: Every 1,000th person on a complete list of all eligible voters in California

Bias. Are there sources of bias in the situations described in Exercises 19–22? Explain.

19. **Movie Critic.** A film critic for ABC News gives her opinion of the latest movie from Disney, which also happens to own ABC.

20. **Car Reviews.** *Consumer Reports* magazine prints a review of new cars and does not accept free products or advertising from anyone.

21. **GMO Soybeans.** Monsanto hires independent university scientists to determine whether its new, genetically engineered soybean poses any threat to the environment.

22. **Drug Study Funding.** The *Journal of the American Medical Association* prints an article evaluating a drug, and some of the physicians who wrote the article received funding from the pharmaceutical company that produces the drug.

Sampling Methods. In Exercises 23–38, identify which of the following applies: simple random sample, systematic sample, convenience sample, stratified sample, or cluster sample. In each case, state whether you think the procedure is likely to yield a representative sample or a biased sample, and explain why.

23. **Clinical Trial.** In phase II testing of a new drug designed to increase the red blood cell count, a researcher obtains envelopes with the names and addresses of all treated subjects. She wants to increase the dosage in a sub-sample of 12 subjects, so she thoroughly mixes all of the envelopes in a bin, then pulls 12 of those envelopes to identify the subjects to be given the increased dosage.

24. **Sobriety Checkpoint.** Police set up a sobriety checkpoint at which every fifth driver is stopped and interviewed.

25. **Exit Polls.** On days of presidential elections, the news media organize an exit poll in which specific polling stations are randomly selected and all voters are surveyed as they leave the premises.

26. **Education and Sports.** A researcher for the Spaulding athletic equipment company is studying the relationship between level of education and participation sports. She conducts a survey of 40 randomly selected golfers, 40 randomly selected tennis players, and 40 randomly selected swimmers.

27. **Ergonomics.** An engineering student measures the strength of fingers used to push buttons by testing family members.

28. **Tax Cheating.** An Internal Revenue Service researcher investigates false reporting of tip income by waiters and waitresses by surveying all waiters and waitresses at 20 randomly selected restaurants.

29. **MTV Survey.** A marketing expert for MTV is planning a survey in which 500 people will be randomly selected from each age group of 10–19, 20–29, and so on.

30. **Credit Card Data.** A professor surveyed students in her class to obtain sample data consisting of the number of credit cards students possess.

31. **Fundraising.** Fundraisers for the College of Newport test a new telemarketing campaign by obtaining an alphabetical list of all alumni and selecting every 100th name on that list.

32. **Telephone Poll.** In a Gallup poll of 1,059 adults, the interview subjects were selected by using a computer to randomly generate telephone numbers that were then called.

33. **Market Research.** A market researcher has partitioned all California residents into categories of unemployed, employed full time, and employed part time. She is surveying 50 people from each category.

34. **Student Drinking.** Motivated by a student who died from binge drinking, the College of Newport conducts a study of student drinking by randomly selecting 10 different classes and interviewing all of the students in each of those classes.

35. **Magazine Survey.** *People* magazine chooses its "best-dressed celebrities" by compiling responses from readers who mail in a survey printed in the magazine.

36. **Heart Transplants.** A medical researcher at Johns Hopkins University obtains a numbered list of all patients waiting for a heart transplant, then uses a computer to select the patients corresponding to the 50 numbers randomly generated by computer.

37. **Quality Control.** A sample of manufactured CDs is obtained by using a computer to randomly generate a number between 1 and 1,000 for each CD, and the CD is selected if the generated number is 1,000.

38. **Seat Belts.** Every 500th seat belt is tested by stressing it until it fails.

Choose a Sampling Method. For each of Exercises 39–42, suggest a sampling method that is likely to produce a representative sample. Explain why you chose this method over other methods.

39. **Student Election.** You want to predict the winner of an upcoming election for student body president.

40. **Blood Type.** You want to determine the percentage of people in this country in each of the four major blood groups (A, B, AB, and O).

41. **Heart Deaths.** You want to determine the percentage of deaths due to heart disease each year.

42. **Mercury in Tuna.** You want to determine the average mercury content of the tuna fish consumed by U.S. residents.

PROJECTS FOR THE INTERNET & BEYOND

43. **Public Opinion Poll.** Use information available on the Web site of a polling organization, such as Gallup, Harris, Pew, or Yankelovich, to answer the following questions.

    a. How exactly is a sample of subjects selected?

    b. Based on what you have learned, do you think the poll results are reliable? If so, why? If not, why not?

44. **Unemployment Sample.** Use the Bureau of Labor Statistics Web page to find details on how the bureau chooses the sample of households in its monthly survey. Write a short summary of the procedure and why it is likely to yield a representative sample.

45. **Selective Voting.** The Academy Awards, the Heisman Trophy, and the *New York Times* "Bestseller List" are just three examples of selections that are determined by the votes of specially selected individuals. Pick one of these selection processes, and describe who votes and how those people are chosen. Discuss sources of bias in the process.

IN THE NEWS

46. **Sampling in the News.** Find a recent news report about a statistical study that you find interesting. Write a short summary of how the sample for the study was chosen, and briefly discuss whether you think the sample was representative of the population under study.

47. **Opinion Poll Sample.** Find a recent news report about an opinion poll carried out by a news organization (such as Gallup, Harris, *USA Today*, *New York Times*, or CNN). Briefly describe the sample and how it was chosen. Was the sample chosen in a way that was likely to introduce any bias? Explain.

48. **Political Polls.** Find results from a recent poll conducted by a political organization (such as the Republican or Democratic party or an organization that seeks to influence Congress on some particular issue). Briefly describe the sample and how it was chosen. Was the sample chosen in a way that was likely to introduce any bias? Should you be more concerned about bias in such a poll than you would be in a poll conducted by a news organization? Explain.

3 TYPES OF STATISTICAL STUDIES

Statistical studies are conducted in many different ways. In all cases, the people, animals (or other living things), or objects chosen for the sample are called the **subjects** of the study. If the subjects are people, it is common to refer to them as **participants** in the study.

Definition

The **subjects** of a study are the people, animals (or other living things), or objects chosen for the sample; if the subjects are people, they may be called the **participants** in the study.

There are two basic types of statistical study: observational studies and experiments. In an **observational study,** we observe or measure specific characteristics while trying to be careful to avoid influencing or modifying the characteristics we are observing. The Nielsen ratings are an example of an observational study, because Nielsen uses devices to *observe* what the subjects are watching on TV, but does not try to influence what they watch.

You can observe a lot by just watching.

—Yogi Berra

Note that an observational study may involve activities that go beyond the usual definition of *observing*. Measuring people's weights requires interacting with them, as in asking them to stand on a scale. But in statistics, we consider these measurements to be observations because the interactions do not change people's weights. Similarly, an opinion poll in which researchers conduct in-depth interviews is considered observational as long as the researchers attempt only to learn people's opinions, not to change them.

In contrast, consider a medical study designed to test whether large daily doses of vitamin C help prevent colds. To conduct this study, the researchers must ask some people in the sample to take large doses of vitamin C every day. This type of statistical study is called an **experiment.** The purpose of an experiment is to study the effects of some **treatment**—in this case, large daily doses of vitamin C.

Two Basic Types of Statistical Study

- In an **observational study,** researchers observe or measure characteristics of the subjects, but do not attempt to influence or modify these characteristics.
- In an **experiment,** researchers apply some **treatment** and observe its effects on the subjects of the experiment.

EXAMPLE 1 Type of Study

Identify the study as an observational study or an experiment.

a. The Salk polio vaccine study (see the Case Study given previously)

b. A poll in which college students are asked if they commute or live on campus

SOLUTION

a. The Salk vaccine study was an *experiment* because researchers tested a treatment—in this case, the vaccine—to see whether it reduced the incidence of polio.

b. The poll is an *observational study* because it attempts to determine how college students go to their classes, and it does not try to influence how they get there. ● ● ●

Identifying the Variables

Statistical studies—whether observations or experiments—generally are attempts to measure what we call **variables of interest.** The term *variable* refers to an item or quantity that can vary or take on different values, and variables of interest are those we seek to learn about. For

example, variables of interest in the Nielsen studies of television viewing habits include *show being watched* and *number of viewers*. The variable *show being watched* can take on different values such as "Super Bowl" or "*60 Minutes*." The variable *number of viewers* depends on the popularity of a particular show. In essence, the raw data in any statistical study are the different values of the variables of interest.

In cases where we think cause and effect may be involved, we sometimes subdivide the variables of interest into two categories. For example, each person in the study of vitamin C and colds may take a different daily dose of vitamin C and may end up with a different number of colds over some period of time. Because we are trying to learn if vitamin C causes a lower number of colds, we say that *daily dose of vitamin C* is an **explanatory variable**—it may explain or cause a change in the number of colds. Similarly, we say that *number of colds* is a **response variable,** because we expect it to respond to changes in the explanatory variable (the dose of vitamin C).

Definitions

A variable is any item or quantity that can vary or take on different values.

The **variables of interest** in a statistical study are the items or quantities that the study seeks to measure.

When cause and effect may be involved, an **explanatory variable** is a variable that may explain or cause the effect, while a **response variable** is a variable that responds to changes in the explanatory variable.

EXAMPLE 2 Identify the Variables

Identify the variables of interest for each study.

a. The Salk polio vaccine study

b. A poll in which college students are asked if they commute or live on campus

SOLUTION

a. The two variables of interest in the Salk vaccine study are *vaccine* and *polio*. They are variables because they can take on two different values: A child either did or did not get the vaccine and either did or did not contract polio. In this case, because the study seeks to determine whether the vaccine prevents polio, we say that *vaccine* is the explanatory variable (it may explain a change in the incidence of polio) and *polio* the response variable (it is supposed to change in response to the vaccine).

b. The variables of interest are the responses to the question asking how students get to their classes, and the proportions of responses to the possible different choices (such as commute, live on campus, refuse to answer, don't know). (There is no cause and effect involved in this study, so we do not need to decide whether the variable is explanatory or response.)

Observational Studies

The observational studies we have discussed up to this point, such as Nielsen ratings, opinion polls, and determining the mean height of students, are studies in which the data are all generally collected around the same time. Sometimes, however, observational studies look at past data or are designed to look at future data over a long period of time.

A **retrospective study** (also called a *case-control* study) is an observational study that uses data from the past—such as official records or past interviews—to learn about some issue of concern. Retrospective studies are especially valuable in cases where it may be impractical or unethical to perform an experiment. For example, suppose we want to learn how alcohol consumed during pregnancy affects newborn babies. Because it is already known that

consuming alcohol during pregnancy can be harmful, it would be highly unethical to ask pregnant mothers to test the "treatment" of consuming alcohol. However, because many mothers consumed alcohol in past pregnancies (either before the dangers were known or choosing to ignore the dangers), we can do a retrospective study in which we compare children born to those mothers to children born to mothers who did not consume alcohol.

Sometimes, the data we need to reach clear conclusions are not available in past records. In those cases, researchers may set up a **prospective study** (sometimes called a *longitudinal* study) designed to collect observations in the future from groups that share common factors. A classic example of a prospective study is the Harvard Nurses' Health Study, which was started in 1976 in order to collect data about how different lifestyles affect women's health (see the "Focus on Public Health" section at the end of this chapter). The study, still ongoing today, has followed thousands of nurses over more than three decades, collecting data about their lifestyles and health.

Variations on Observational Studies

The most familiar observational studies are those in which data are collected all at once (or as close to that as possible). Two variations on observational studies are also common:

- A **retrospective** (or **case-control**) **study** uses data from the past, such as official records or past interviews.
- A **prospective** (or **longitudinal**) **study** is set up to collect data in the future from groups that share common factors.

EXAMPLE 3 Observational Study

You want to know whether children born prematurely do as well in elementary school as children born at full term. What type of study should you do?

SOLUTION An observational, retrospective study is the only real option in this case. You would collect data on past births and compare the elementary school performance of those born prematurely to that of those born at full term.

Experiments

Because experiments require active intervention, such as applying a treatment, we must take special care to ensure that they are designed in ways that will provide the information we seek. Let's examine a few of the issues that arise in the design of experiments.

The Need for Controls

Consider an experiment that gives some participants in a study vitamin C to determine its effect on colds. Suppose the people taking vitamin C daily get an average of 1.5 colds in a three-month period. How can the researchers know whether the subjects would have gotten more colds without the vitamin C? To answer this type of question, the researchers must conduct their experiment with two (or more) groups of subjects: One group takes large doses of vitamin C daily and another group does not. As we'll discuss shortly, in most cases it is important that participants be *randomly* assigned to the two groups.

The group of people who are randomly assigned to take vitamin C is called the **treatment group** because its members receive the treatment being tested (vitamin C). The group of people who do *not* take vitamin C is called the **control group.** The researchers can be confident that vitamin C is an effective treatment only if the people in the treatment group get significantly fewer colds than the people in the control group.

BY THE WAY

The control group gets its name from the fact that it helps control the way we interpret experimental results.

Treatment and Control Groups

The **treatment group** in an experiment is the group of subjects who receive the treatment being tested.

The **control group** in an experiment is the group of subjects who do *not* receive the treatment being tested.

EXAMPLE 4 Treatment and Control

Look again at the Salk polio vaccine Case Study. What was the treatment? Which group of children constituted the treatment group? Which constituted the control group?

SOLUTION The treatment was the Salk vaccine. The treatment group consisted of the children who received the Salk vaccine. The control group consisted of the children who did not get the Salk vaccine and instead got an injection of salt water. • • ●

EXAMPLE 5 Mozart Treatment

A study divided college students into two groups. One group listened to Mozart or other classical music before being assigned a specific task and the other group simply was assigned the task without listening to the music. Researchers found that those listening to the classical music performed the task slightly better, but only if they did the task within a few minutes of listening to the music. (The two groups performed equally on tasks given later). Identify the treatment and the control and treatment groups.

SOLUTION The treatment was the classical music. The treatment group consisted of the students who listened to the music. The control group consisted of the students who did not listen to the music. • • ●

BY THE WAY

The so-called Mozart effect holds that listening to Mozart can make babies smarter. The supposed effect spawned an entire industry of Mozart products for children. The state of Georgia even began passing out Mozart CDs to new mothers. However, more recent studies of the Mozart effect have been unable to substantiate the claimed effect.

Confounding Variables

Using control groups helps to ensure that we account for known variables that could affect a study's results. However, researchers may be unaware of or be unable to account for other important variables. Consider an experiment in which a statistics teacher seeks to determine whether students who study collaboratively (in study groups with other students) earn higher grades than students who study independently. The teacher chooses five students who will study collaboratively (the treatment group) and five others who will study independently (the control group). To ensure that the students all have similar abilities and will study diligently, the teacher chooses only students with high grade-point averages. At the end of the semester, the teacher finds that the students who studied collaboratively earned higher grades.

The variables of interest for this study are *collaborative study* (whether they do so or not) and *final grade*. But suppose that, unbeknownst to the teacher, the collaborative students all lived in a dormitory where a curfew ensured that they got plenty of sleep. This fact introduces a new variable—which we might call *amount of sleep*—that might partially explain the results. In other words, the experiment's conclusion may *seem* to support the benefits of collaborative study, but this conclusion is not justified because the teacher did not account for how much students slept.

In statistical terminology, this study suffers from **confounding.** The higher grades may be due either to the variable of interest (*collaborative study*) or to the differing amounts of sleep or to a combination of both. Because the teacher did not account for differences in the amount of sleep, we say that *amount of sleep* is a **confounding variable** for this study. You can probably think of other potentially confounding variables that could affect a study like this one.

Definition

A study suffers from **confounding** if the effects of different variables are mixed so we cannot determine the specific effects of the variables of interest. The variables that lead to the confusion are called **confounding variables.**

CASE STUDY Confounding Drug Results

An advisory panel of the Federal Drug Administration (FDA) recently recommended revoking the approval of the advanced breast cancer drug Avastin, which at the time was the world's best-selling cancer drug. Early trials had found that, when used in combination with a certain chemotherapy drug, Avastin delayed the growth of tumors by five months compared to using the chemotherapy drug alone. However, additional trials found less significant delays and no improvement in the length or quality of the lives of those taking Avastin. The differing results in different trials suggest that confounding variables have not all been accounted for. One possible confounding variable may be the choice of chemotherapy drug. Indeed, studies in Europe found Avastin to be effective, but were based on using it in combination with a different chemotherapy drug than the one approved in the United States. As a result, European regulators were approving Avastin for expanded use at the same time that regulators in the United States were recommending against it. The lesson should be clear: Even when thousands of lives and millions of dollars are on the line, confounding variables may be very difficult to eliminate, making decisions far more uncertain than doctors or patients (or government regulators) would like.

Assigning Treatment and Control Groups

As the collaborative study experiment illustrates, results are almost sure to suffer from confounding if the treatment and control groups differ in important ways (other than receiving or not receiving the treatment). Researchers generally employ two strategies to prevent such differences and thereby ensure that the treatment and control groups can be compared fairly. First, they assign participants to the treatment and control groups *at random*, meaning that they use a technique designed to ensure that each participant has an equal chance of being assigned to either group. When the participants are randomly assigned, it is less likely that the people in the treatment and control groups will differ in some way that will affect the study results.

Second, researchers try to ensure that the treatment and control groups are sufficiently large. For example, in the collaborative study experiment, including 50 students in each group rather than five would have made it much less likely that all the students in one group would live in a special dormitory.

Strategies for Selecting Treatment and Control Groups

- **Select groups at random.** Make sure that the subjects of the experiment are assigned to the treatment or control group at random, meaning that each subject has an equal chance of being assigned to either group.
- **Use sufficiently large groups.** Make sure that the treatment and control groups are both sufficiently large that they are unlikely to differ in a significant way (aside from the fact that one group gets the treatment and the other does not).

EXAMPLE 6 Salk Study Groups

Briefly explain how the two strategies for selecting treatment and control groups were used in the Salk polio vaccine study.

SOLUTION A total of about 400,000 children participated in the study, with half receiving an injection of the Salk vaccine (the treatment group) and the other half receiving an injection of salt water (the control group). The first strategy was implemented by choosing children for the two groups randomly from among all the children. The second strategy was implemented by using a large number of participants (200,000 in each group) so that the two groups were unlikely to differ by chance.

There are men on whom the mere sight of medicine is operative.

—French philosopher Michel de Montaigne (1533–1592)

The Placebo Effect

When an experiment involves people, effects can occur simply because people know they are part of the experiment. For example, suppose you are testing the effectiveness of a new anti-depression drug. You find 500 people who suffer from depression and randomly divide them into a treatment group that receives the new drug and a control group that does not. A few weeks later, interviews with the patients show that people in the treatment group tend to be feeling much better than people in the control group. Can you conclude that the new drug works?

Unfortunately, it's quite possible that the mood of people receiving the drug improved simply because they were happy to be getting some kind of treatment, which means you cannot be sure that the drug really helped. This type of effect, in which people improve because they believe that they are receiving a useful treatment, is called the **placebo effect.** (The word *placebo* comes from the Latin "to please.")

To distinguish between results caused by a placebo effect and results that are truly due to the treatment, researchers try to make sure that the participants do not know whether they are part of the treatment or control group. To accomplish this, the researchers give the people in the control group a **placebo**: something that looks or feels just like the treatment being tested, but lacks its active ingredients. For example, in a test of a drug that comes in pill form, the placebo might be a pill of the same shape and size that contains sugar instead of the real drug. In a test of an injected vaccine, the placebo might be an injection that contains only a saline solution (salt water) instead of the real vaccine. In a recent test of the effectiveness of acupuncture, the placebo consisted of treatment with needles as in real acupuncture, except the needles were not put in the special places that acupuncturists claim to be important.

As long as the participants do not know whether they received the real treatment or a placebo, the placebo effect ought to affect the treatment and control groups equally. If the results for the two groups are significantly different, it is reasonable to believe that the differences can be attributed to the treatment. For example, in the study of the anti-depression drug, we would conclude that the drug was effective only if the control group received a placebo and members of the treatment group improved much more than members of the control group. For even better control, some experiments use three groups: a treatment group, a placebo group, and a control group. The placebo group is given a placebo while the control group is given nothing.

BY THE WAY

The placebo effect can be remarkably powerful. In some studies, up to 75% of the participants receiving the placebo actually improve. Nevertheless, different researchers disagree about the strength of the placebo effect, and some even question the reality of the effect.

Definitions

A placebo lacks the active ingredients of a treatment being tested in a study, but looks or feels like the treatment so that participants cannot distinguish whether they are receiving the placebo or the real treatment.

The **placebo effect** refers to the situation in which patients improve simply because they believe they are receiving a useful treatment.

Note: Although participants should not know whether they belong to the treatment or control group, for ethical reasons it is very important that participants be told that some of them will be given a placebo, rather than the real treatment.

BY THE WAY

A related effect, known as the *Hawthorne effect*, occurs when treated subjects somehow respond differently simply because they are part of an experiment—regardless of the particular way in which they are treated. The Hawthorne effect gets its name from the fact that it was first observed in a study of factory workers at Western Electric's Hawthorne plant.

EXAMPLE 7 Vaccine Placebo

What was the placebo in the Salk polio vaccine study? Why did researchers use a placebo in this experiment?

SOLUTION The placebo was the salt water injection given to the children in the control group. To understand why the researchers used a placebo for the control group, suppose that a placebo had *not* been used. When improvements were observed in the treatment group, it would have been impossible to know whether the improvements were due to the vaccine or to the placebo effect. In order to remove this confounding, all participants had to believe that they were being treated in the same way. This ensured that any placebo effect would occur in both groups equally, so that researchers could attribute any remaining differences to the vaccine.

TIME OUT TO THINK

Although participants should not know whether they belong to the treatment or control (placebo) group, for ethical reasons they should be told that some of them will receive a placebo rather than the real treatment. This was not always the case in decades past, when participants were sometimes told that they all received a treatment when in fact some received a placebo. Should researchers be allowed to use results of past studies that do not meet today's ethical criteria? Defend your opinion.

Experimenter Effects

Even if the study subjects don't know whether they received the real treatment or a placebo, the *experimenters* may still have an effect. In testing an anti-depression drug, for example, experimenters will probably interview patients to find out whether they are feeling better. But if the experimenters know who received the real drug and who received the placebo, they may inadvertently smile more at the people in the treatment group. Their smiles might improve those participants' moods, making it seem as if the treatment worked when in fact the improvement was caused by the experimenter. This type of confounding, in which the experimenter somehow influences the results, is called an **experimenter effect** (or a Rosenthal effect). The only way to avoid experimenter effects is to make sure that the experimenters don't know which subjects are in which group.

Definition

An **experimenter effect** occurs when a researcher or experimenter somehow influences subjects through such factors such as facial expression, tone of voice, or attitude.

EXAMPLE 8 Child Abuse?

In a famous case, two couples from Bakersfield, California, were convicted of molesting dozens of preschool-age children at their daycare center. The evidence for the abuse came primarily from interviews with the children. However, the conviction was overturned—after one man had served 14 years in prison—when a judge re-examined the interviews and concluded that the children had given answers that they thought the interviewers wanted to hear. If we think of the interviewers as experimenters, this is an example of an experimenter effect because the interviewers influenced the children's answers through the tone and style of their questioning.

Blinding

In statistical terminology, the practice of keeping people in the dark about who is in the treatment group and who is in the control group is called **blinding.** A **single-blind** experiment is one in which the participants don't know which group they belong to, but the experimenters do know. If neither the participants nor the experimenters know who belongs to each group, the study is said to be **double-blind.** Of course, *someone* has to keep track of the two groups in order to evaluate the results at the end. In a double-blind experiment, the researchers conducting the study typically hire experimenters to make any necessary contact with the participants. The researchers thereby avoid any contact with the participants, ensuring that they cannot influence them in any way. The Salk polio vaccine study was double-blind because neither the participants (the children) nor the experimenters (the doctors and nurses giving the injections and diagnosing polio) knew who got the real vaccine and who got the placebo.

Blinding in Experiments

An experiment is **single-blind** if the participants do not know whether they are members of the treatment group or members of the control group, but the experimenters do know.

An experiment is **double-blind** if neither the participants nor any experimenters know who belongs to the treatment group and who belongs to the control group.

BY THE WAY

Many similar cases of supposedly widespread child abuse at daycare centers and preschools are being re-examined to see if experimenter effects (by those who interviewed the children) may have led to wrongful convictions. Similar claims of experimenter effects have been made in cases involving repressed memory, in which counseling supposedly helped people retrieve lost memories of traumatic events.