CHAPTER 7

The Logic of Sampling

CHAPTER OVERVIEW

Now you'll see how social scientists can select a few people for study—and discover things that apply to hundreds of millions of people not studied.

iStock.com/Gary Blakeley

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Learning Objectives

After studying this chapter, you will be able to . . .

- Highlight some of the key events in the development of sampling in social research.
- Describe what is meant by "nonprobability sampling" and identify several techniques.
- Identify and explain the key elements in probability sampling.
- Explain the relationship between populations and sampling frames in social research.
- Identify and describe several types of probability sampling designs.
- Describe the steps involved in selecting a multistage cluster sample.
- Discuss the key advantages of probability sampling.
- Explain how the sampling design of a study could have ethical implications.

Introduction

One of the most visible uses of survey sampling lies in the political polling that is subsequently tested by election results. Whereas some people doubt the accuracy of sample surveys, others complain that political polls take all the suspense out of campaigns by foretelling the result.

Going into the 2008 presidential elections, pollsters were in agreement as to who would win, in contrast to their experiences in 2000 and 2004, which were closely contested races. Table 7-1 reports polls conducted during the few days preceding the election. Despite some variations, the overall picture they present is amazingly consistent and pretty well matches the election results.

Now, how many interviews do you suppose it took each of these pollsters to come within a couple of percentage points in estimating the behavior of more than 131 million voters? Often fewer than 2,000! In this chapter, we're going to find out how social researchers can achieve such wizardry.

In the 2016 presidential election, the pre-election polls again clustered closely around the actual popular votes for Hillary Clinton and Donald Trump. Most correctly predicted that Secretary Clinton would win the popular vote by 2 or 3 percentage points.

Of course, the president is not elected by the nation's overall popular vote, but by the electoral college, determined by how the votes go in the individual states. Relatively small victories totaling 107,000 votes in three swing states—Michigan, Pennsylvania, Wisconsin—gave Trump all those states' electoral votes, and the presidency, while Clinton won the popular vote by 2.8 million (Washington Post 2016).

FiveThirtyEight.com offers a useful analysis and rating of the many polling companies active in forecasting political outcomes.

TABLE 7-1

Election-Eve Polls Reporting Presidential Voting Plans, 2008

Poll	*Date Ended*	*Obama*	*McCain*
Fox	Nov 2	54	46
NBC/WSJ	Nov 2	54	46
Marist College	Nov 2	55	45
Harris Interactive	Nov 3	54	46
Reuters/C-SPAN/Zogby	Nov 3	56	44
ARG	Nov 3	54	46
Rasmussen	Nov 3	53	47
IBD/TIPP	Nov 3	54	46
DailyKos.com/Research 2000	Nov 3	53	47
GWU	Nov 3	53	47
Marist College	Nov 3	55	45
Actual vote	Nov 4	54	46

Note: For simplicity, since there were no "undecideds" in the official results and each of the third-party candidates received less than one percentage of the vote, I've apportioned the undecided and other votes according to the percentages saying they were voting for Obama or McCain.

Source: Poll data are adapted from http://www.pollster.com/polls/us/08-us-pres-ge-mvo.php. The official election results are from the Federal Election Commission, http://www.fec.gov/pubrec/fe2008/2008presgeresults.pdf.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

What do you think?

In 1936, the *Literary Digest* collected the voting intentions of 2 million voters in order to predict whether Franklin D. Roosevelt or Alf Landon would be elected president of the United States. During more-recent election campaigns, with many more voters going to the polls, national polling firms have typically sampled around 2,000 voters across the country.

Which technique do you think is the most effective? Why?

See the *What do you think? ... Revisited* box toward the end of the chapter.

© Christopher Halloran/Shutterstock.com

For another powerful illustration of the potency of sampling, look at Figure 7-1 for a graph of then-president George W. Bush's approval ratings prior to and following the September 11, 2001, terrorist attacks on the United States. The data reported by several different polling agencies describe the same pattern.

Political polling, like other forms of social research, rests on observations. But neither pollsters nor other social researchers can observe everything that might be relevant to their interests. A critical part of social research, then, is deciding what to observe and what not to observe. If you want to study voters, for example, which voters should you study?

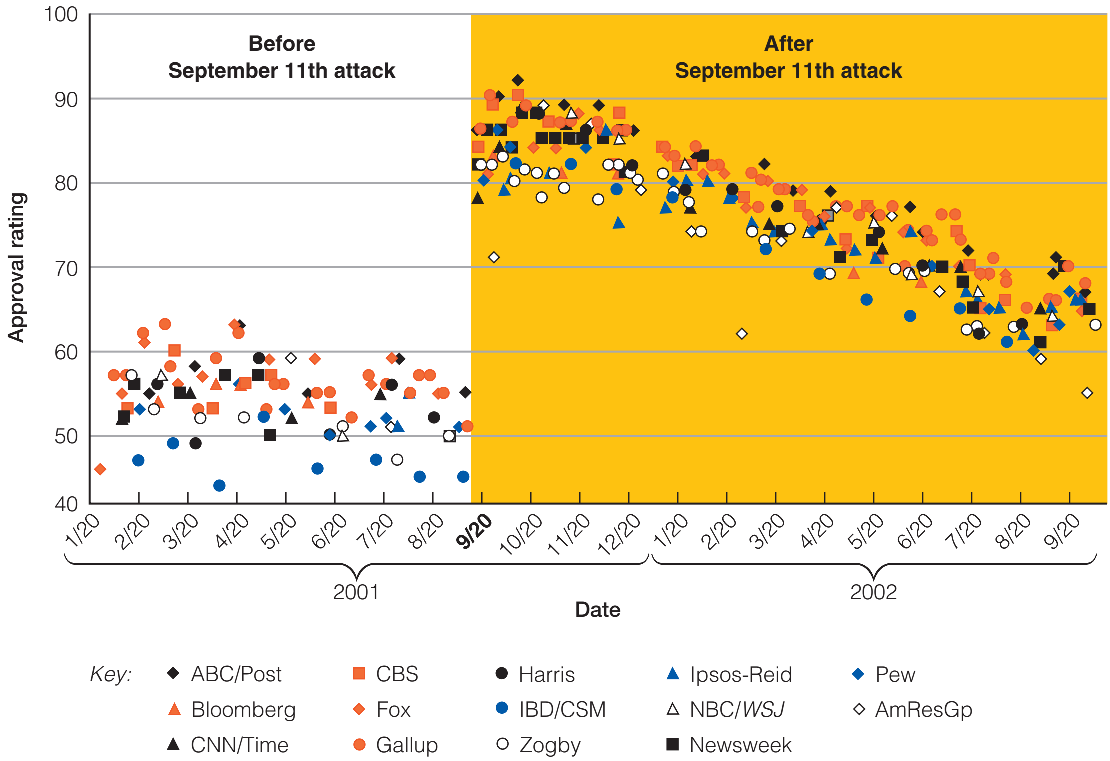


FIGURE 7-1

Bush Approval: Raw Poll Data. This graph demonstrates how independent polls produce the same picture of reality. It also shows the impact of a national crisis on the president's popularity: in this case, the 9/11 terrorist attack and then-president George W. Bush's popularity.

Source: drlimerick.com

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

The process of selecting observations is called *sampling*. Although sampling can mean any procedure for selecting units of observation—for example, interviewing every tenth passerby on a busy street—the key to generalizing from a sample to a larger population is probability sampling, which involves the important idea of random selection.

Much of this chapter is devoted to the logic and skills of probability sampling. This topic is more rigorous and precise than some of the other topics in this book. Whereas social research as a whole is both art and science, sampling leans toward science. Although this subject is somewhat technical, the basic logic of sampling is not difficult to understand. In fact, the logical neatness of this topic can make it easier to comprehend than, say, conceptualization.

Although probability sampling is central to social research today, we'll also examine a variety of nonprobability methods. These methods have their own logic and can provide useful samples for social inquiry.

Before we discuss the two major types of sampling, I'll introduce you to some basic ideas by way of a brief history of sampling. As you'll see, the pollsters who correctly predicted recent elections have done so in part because researchers had learned to avoid some pitfalls that earlier pollsters had discovered "the hard way."

A Brief History of Sampling

Sampling in social research has developed hand in hand with political polling. This is the case, no doubt, because political polling is one of the few opportunities social researchers have to discover the accuracy of their estimates. On election day, they find out how well or how poorly they did.

President Alf Landon

President Alf Landon? Who's he? Did you sleep through an entire presidency in your U.S. history class? No—but Alf Landon would have been president if a famous poll conducted by the *Literary Digest* had proved to be accurate. The *Literary Digest* was a popular newsmagazine published between 1890 and 1938. In 1916, *Digest* editors mailed postcards to people in six states, asking them whom they were planning to vote for in the presidential campaign between Woodrow Wilson and Charles Evans Hughes. Names were selected for the poll from telephone directories and automobile registration lists. Based on the postcards sent back, the *Digest* correctly predicted that Wilson would be elected. In the elections that followed, the *Literary Digest* expanded the size of its poll and made correct predictions in 1920, 1924, 1928, and 1932.

In 1936 the *Digest* conducted its most ambitious poll: 10 million ballots were sent to people listed in telephone directories and on lists of automobile owners. Over 2 million people responded, giving the Republican contender, Alf Landon, a stunning 57 to 43 percent landslide over the incumbent, President Franklin Roosevelt. The editors modestly cautioned,

> *We make no claim to infallibility. We did not coin the phrase "uncanny accuracy" which has been so freely applied to our Polls. We know only too well the limitations of every straw vote, however enormous the sample gathered, however scientific the method. It would be a miracle if every State of the forty-eight behaved on Election Day exactly as forecast by the Poll.*
>
> (Literary Digest *1936a: 6*)

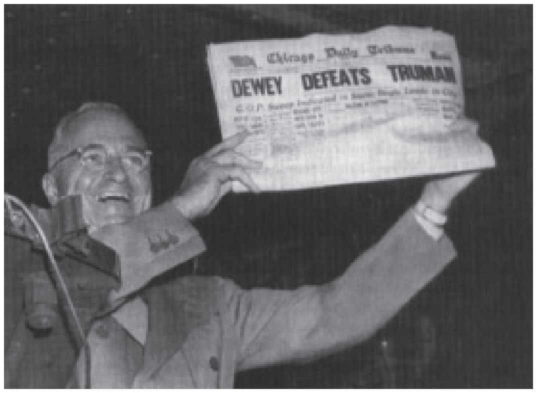
Two weeks later, the *Digest* editors knew the limitations of straw polls even better: The voters gave Roosevelt a second term in office by the largest landslide in history, with 61 percent of the vote. Landon won only 8 electoral votes to Roosevelt's 523.

The editors were puzzled by their unfortunate turn of luck. Part of the problem surely lay in the 22 percent return rate garnered by the poll. The editors asked,

> *Why did only one in five voters in Chicago to whom the* Digest *sent ballots take the trouble to reply? And why was there a preponderance of Republicans in the one-fifth that did reply?. . . We were getting better cooperation in what we have always regarded as a public service from Republicans than we were getting from Democrats. Do Republicans live nearer to mailboxes? Do Democrats generally disapprove of straw polls?*
>
> (Literary Digest *1936b: 7*)

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Actually, there was a better explanation—and it lay in what is technically called the *sampling frame* used by the *Digest*. In this case the sampling frame consisted of telephone subscribers and automobile owners. In the context of 1936, this design selected a disproportionately wealthy sample of the voting population, especially coming on the tail end of the worst economic depression in the nation's history. The sample effectively excluded poor people, and the poor voted predominantly for Roosevelt's New Deal recovery program. The *Digest*'s poll may or may not have correctly represented the voting intentions of telephone subscribers and automobile owners. Unfortunately for the editors, it decidedly did not represent the voting intentions of the population as a whole.



Bettmann/Getty Images

Basing its decision on early political polls that showed Dewey leading Truman, the *Chicago Tribune* sought to scoop the competition with this unfortunate headline.

President Thomas E. Dewey

The 1936 election also saw the emergence of a young pollster whose name would become synonymous with public opinion. In contrast to the *Literary Digest*, George Gallup correctly predicted that Roosevelt would beat Landon. Gallup's success in 1936 hinged on his use of something called *quota sampling*, which we'll examine later in the chapter. For now, it's enough to know that quota sampling is based on a knowledge of the characteristics of the population being sampled: the proportion of men, the proportion of women, that proportions of various incomes, ages, and so on. Quota sampling selects people to match a set of these characteristics: the right number of poor, white, rural men; the right number of rich, African American, urban women; and so on. The quotas are based on those variables most relevant to the study. In the case of Gallup's poll, the sample selection was based on levels of income; the selection procedure ensured the right proportion of respondents at each income level.

Gallup and his American Institute of Public Opinion used quota sampling to good effect in 1936, 1940, and 1944—correctly picking the presidential winner each time. Then, in 1948, Gallup and most political pollsters suffered the embarrassment of picking Governor Thomas Dewey of New York over the incumbent, President Harry Truman. The pollsters' miscue continued right up to election night. A famous photograph shows a jubilant Truman—whose followers' battle cry was "Give 'em hell, Harry!"—holding aloft a newspaper with the banner headline "Dewey Defeats Truman."

Several factors accounted for the pollsters' failure in 1948. First, most pollsters stopped polling in early October, despite a steady trend toward Truman toward the end of the campaign. In addition, many voters were undecided throughout the campaign, and they went disproportionately for Truman when they stepped into the voting booth.

More important, Gallup's failure rested on the unrepresentativeness of his samples. Quota sampling—which had been effective in earlier years—was Gallup's undoing in 1948. This technique requires that the researcher know something about the total population (of voters, in this instance). For national political polls, such information came primarily from census data. By 1948, however, World War II had produced a massive movement from the country to cities, radically changing the character of the U.S. population from what the 1940 census showed, and Gallup relied on 1940 census data. City dwellers, moreover, tended to vote Democratic; hence, the overrepresentation of rural voters in his poll had the effect of underestimating the number of Democratic votes.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Two Types of Sampling Methods

By 1948 some academic researchers had already been experimenting with a form of sampling based on probability theory. This technique involves the selection of a "random sample" from a list containing the names of everyone in the population being sampled. By and large, the probability-sampling methods used in 1948 were far more accurate than quota-sampling techniques.

Today, probability sampling remains the primary method of selecting large, representative samples for social research, including national political polls. At the same time, probability sampling can be impossible or inappropriate in many research situations. Accordingly, before turning to the logic and techniques of probability sampling, we'll first take a look at techniques for nonprobability sampling and how they're used in social research.

Nonprobability Sampling

Social research is often conducted in situations that do not permit the kinds of probability samples used in large-scale social surveys. Suppose you wanted to study homelessness: There is no list of all homeless individuals, nor are you likely to create such a list. Moreover, as you'll see, there are times when probability sampling would not be appropriate even if it were possible. Many such situations call for **nonprobability sampling**.

In this section, we'll examine four types of nonprobability sampling: reliance on available subjects, purposive or judgmental sampling, snowball sampling, and quota sampling. We'll conclude with a brief discussion of techniques for obtaining information about social groups through the use of informants.

nonprobability sampling Any technique in which samples are selected in some way not suggested by probability theory. Examples include reliance on available subjects as well as purposive (judgmental), snowball, and quota sampling.

Reliance on Available Subjects

Relying on available subjects, such as stopping people at a street corner or some other location, is sometimes called "convenience" or "haphazard" sampling. This is a common method for journalists in their "person-on-the-street" interviews, but it is an extremely risky sampling method for social research. Clearly, this method does not permit any control over the representativeness of a sample. It's justified only if the researcher wants to study the characteristics of people passing the sampling point at specified times or if less risky sampling methods are not feasible. Even when this method is justified on grounds of feasibility, researchers must exercise great caution in generalizing from their data. Also, they should alert readers to the risks associated with this method.

University researchers frequently conduct surveys among the students enrolled in large lecture classes. The ease and frugality of this method explains its popularity, but it seldom produces data of any general value. It may be useful for pretesting a questionnaire, but such a sampling method should not be used for a study purportedly describing the student body as a whole.

Consider this report on the sampling design in an examination of knowledge and opinions about nutrition and cancer among medical students and family physicians:

> *The fourth-year medical students of the University of Minnesota Medical School in Minneapolis comprised the student population in this study. The physician population consisted of all physicians attending a "Family Practice Review and Update" course sponsored by the University of Minnesota Department of Continuing Medical Education.*
>
> *(Cooper-Stephenson and Theologides 1981: 472)*

After all is said and done, what will the results of this study represent? They do not provide a meaningful comparison of medical students and family physicians in the United States or even in Minnesota. Who were the physicians who attended the course? We can guess that they were probably more concerned about their continuing education than were other physicians, but we can't say for sure.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Although such studies can provide useful insights, we must take care not to overgeneralize from them.

Purposive or Judgmental Sampling

Sometimes it's appropriate to select a sample on the basis of knowledge of a population, its elements, and the purpose of the study. This type of sampling is called **purposive sampling** (or *judgmental sampling*). In the initial design of a questionnaire, for example, you might wish to select the widest variety of respondents to test the broad applicability of questions. Although the study findings would not represent any meaningful population, the test run might effectively uncover any peculiar defects in your questionnaire. This situation would be considered a pretest, however, rather than a final study.

In some instances, you may wish to study a small subset of a larger population in which many members of the subset are easily identified, but the enumeration of them all would be nearly impossible. For example, you might want to study the leadership of a student-protest movement; many of the leaders are visible, but it would not be feasible to define and sample all leaders. In studying all or a sample of the most visible leaders, you may collect data sufficient for your purposes.

Or let's say you want to compare left-wing and right-wing students. Because you may not be able to enumerate and sample from all such students, you might decide to sample the memberships of left- and right-leaning groups, such as the Green Party and the Young Americans for Freedom. Although such a sample design would not provide a good description of either left-wing or right-wing students as a whole, it might suffice for general comparative purposes.

Field researchers are often particularly interested in studying deviant cases—those do not fit into patterns of mainstream attitudes and behaviors—in order to improve their understanding of the more usual pattern. For example, you might gain important insights into the nature of school spirit, as exhibited at a pep rally, by interviewing people who did not appear to be caught up in the emotions of the crowd or by interviewing students who did not attend the rally at all. Selecting deviant cases for study is another example of purposive study.

In qualitative research projects, the sampling of subjects may evolve as the structure of the situation being studied becomes clearer and certain types of subjects seem more central to understanding than others. Let's say you're conducting an interview study among the members of a radical political group on campus. You may initially focus on friendship networks as a vehicle for the spread of group membership and participation. In the course of your analysis of the earlier interviews, you may find several references to interactions with faculty members in one of the social science departments. As a consequence, you may expand your sample to include faculty in that department and other students that they interact with. This is called "theoretical sampling," since the evolving theoretical understanding of the subject steers the sampling in certain directions.

Snowball Sampling

Another nonprobability-sampling technique, which some consider to be a form of accidental sampling, is called **snowball sampling**. This procedure is appropriate when the members of a special population are difficult to locate, such as homeless individuals, migrant workers, or undocumented immigrants. In snowball sampling, the researcher collects data on the few members of the target population he or she can locate, then asks those individuals to provide the information needed to locate other members of that population whom they happen to know. "Snowball" refers to the process of accumulation as each located subject suggests other subjects. Because this procedure also results in samples with questionable representativeness, it's used primarily for exploratory purposes. Sometimes, the term *chain referral* is used in reference to

purposive sampling A type of nonprobability sampling in which the units to be observed are selected on the basis of the researcher's judgment about which ones will be the most useful or representative. Also called *judgmental sampling*.

snowball sampling A nonprobability-sampling method, often employed in field research, whereby each person interviewed may be asked to suggest additional people for interviewing.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

snowball sampling and other, similar techniques in which the sample unfolds and grows from an initial selection.

Suppose you wish to learn a community organization's pattern of recruitment over time. You might begin by interviewing fairly recent recruits, then asking them who introduced them to the group. You might then interview the people named, asking *them* who introduced them to the group. You might then interview the next round of people named, and so forth. Or, in studying a loosely structured political group, you might ask one of the participants who he or she believes to be the most influential members of the group. You might interview those people and, in the course of the interviews, ask who they believe to be the most influential. In each of these examples, your sample would "snowball" as each of your interviewees suggested other people to interview.

In another example, Karen Farquharson (2005) provides a detailed discussion of how she used snowball sampling to discover a network of tobacco policy makers in Australia: both those at the core of the network and those on the periphery.

Kath Browne (2005) used snowballing through social networks to develop a sample of nonheterosexual women in a small town in the United Kingdom. She reports that her own membership in such networks greatly facilitated this type of sampling and that potential subjects in the study were more likely to trust her than to trust heterosexual researchers.

In more general, theoretical terms, Chaim Noy argues that the process of selecting a snowball sample reveals important aspects of the populations being sampled: "the dynamics of natural and organic social networks" (2008: 329). Do the people you interview know others like themselves? Are they willing to identify those people to researchers? In this way, snowball sampling can be more than a simple technique for finding people to study. It, in itself, can be a revealing part of the inquiry.

quota sampling A type of nonprobability sampling in which units are selected for a sample on the basis of prespecified characteristics, so that the total sample will have the same distribution of characteristics assumed to exist in the population being studied.

Jaime Waters (2015) discovered some of the limitations of snowball sampling. In an attempt to study adult (over 40) users of illegal drugs, he discovered that his initial subjects were reluctant or unable to identify other users. Partly, this seemed to reflect a feeling that they had more to lose if their drug use were discovered. Also, he found that his adult users were not as involved in drug-using networks as younger users. Still, snowball sampling is sometimes an effective way to reach hard-to-find subjects.

Quota Sampling

Quota sampling is the method that helped George Gallup avoid disaster in 1936—and set up the disaster of 1948. Like probability sampling, quota sampling addresses the issue of representativeness, although the two methods approach the issue quite differently.

Quota sampling begins with a matrix, or table, describing the characteristics of the target population. Depending on your research purposes, you may need to know what proportion of the population is male and what proportion female, as well as what proportions of each sex fall into various categories of age, educational level, ethnic group, and so forth. In establishing a national quota sample, you might need to know what proportion of the national population is urban, Eastern, male, under 25, white, working class, and the like, and all the possible combinations of these attributes.

Once you've created such a matrix and assigned a relative proportion to each cell in the matrix, you proceed to collect data from people having all the characteristics of a given cell. You then assign to all the people in a given cell a weight appropriate to their portion of the total population. When all the sample elements are so weighted, the overall data should provide a reasonable representation of the total population.

Although quota sampling resembles probability sampling, it has several inherent problems. First, the quota frame (the proportions that different cells represent) must be accurate, and it's often difficult to get up-to-date information for this purpose. The Gallup failure

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

to predict Truman as the presidential victor in 1948 stemmed partly from this problem. Second, the selection of sample elements within a given cell may be biased even if its proportion of the population is accurately estimated. Instructed to interview five people who meet a given, complex set of characteristics, an interviewer may still avoid people living at the top of seven-story walk-ups, having particularly run-down homes, or owning vicious dogs.

In recent years, some researchers have attempted to combine probability and quota-sampling methods, but the effectiveness of this effort remains to be seen. At present, you should treat quota sampling warily if your purpose is statistical description.

At the same time, the logic of quota sampling can sometimes be applied usefully to a field research project. In the study of a formal group, for example, you might wish to interview both leaders and nonleaders. In studying a student political organization, you might want to interview radical, moderate, and conservative members of that group. You may be able to achieve sufficient representativeness in such cases by using quota sampling to ensure that you interview both men and women, both younger and older people, and so forth.

J. Michael Brick (2011), in pondering the future of survey sampling, suggested the possibility of a rebirth for quota sampling. Perhaps it is a workable solution to the problem of representativeness that bedevils falling response rates and online surveys. We'll return to this issue in Chapter 9 on survey research.

Selecting Informants

When field research involves the researcher's attempt to understand some social setting—a juvenile gang or local neighborhood, for example—much of that understanding will come from a collaboration with some members of the group being studied. Whereas social researchers speak of *respondents* as people who provide information about themselves, allowing the researcher to construct a composite picture of the group those respondents represent, an **informant** is a member of the group who can talk directly about the group per se.

Anthropologists in particular depend on informants, but other social researchers rely on them as well. If you wanted to learn about informal social networks in a local public-housing project, for example, you would do well to locate individuals who understand what you are looking for and help you find it.

When Jeffrey Johnson (1990) set out to study a salmon-fishing community in North Carolina, he used several criteria to evaluate potential informants. Did their positions allow them to interact regularly with other members of the camp, for example, or were they isolated? (He found that the carpenter had a wider range of interactions than did the boat captain.) Was their information about the camp limited to their specific jobs, or did it cover many aspects of the operation? These and other criteria helped determine how useful the potential informants might be to his study. We'll return to this example in a bit.

Usually, you'll want to select informants who are somewhat typical of the groups you're studying. Otherwise, their observations and opinions may be misleading. Interviewing only physicians will not give you a well-rounded view of how a community medical clinic is working, for example. Along the same lines, an anthropologist who interviews only men in a society where women are sheltered from outsiders will get a biased view. Similarly, although informants fluent in English are convenient for English-speaking researchers from the United States, they do not typify the members of many societies or even many subgroups within English-speaking countries.

Simply because they're the ones willing to work with outside investigators, informants will almost always be somewhat "marginal" or atypical within their group. Sometimes this is obvious. Other times, however, you'll learn about their marginality only in the course of your research.

In Johnson's study, a county agent identified one fisherman who seemed squarely in the mainstream of the community. Moreover, he was cooperative and helpful to Johnson's research. The more

informant Someone who is well versed in the social phenomenon that you wish to study and who is willing to tell you what he or she knows about it. Not to be confused with a respondent.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Earl Babbie

With so many possible informants, how can the researcher begin to choose?

Johnson worked with the fisherman, however, the more he found the man to be a marginal member of the fishing community.

> *First, he was a Yankee in a southern town. Second, he had a pension from the Navy [so he was not seen as a "serious fisherman" by others in the community]. . . . Third, he was a major Republican activist in a mostly Democratic village. Finally, he kept his boat in an isolated anchorage, far from the community harbor.*
>
> *(Johnson 1990: 56)*

Informants' marginality may not only bias the view you get but also limit their access (and hence yours) to the different sectors of the community you wish to study.

These comments should give you some sense of the concerns involved in nonprobability sampling, typically used in qualitative research projects. I conclude with the following injunction from John Lofland, a particularly thoughtful and experienced qualitative researcher:

> *Your overall goal is to collect the richest possible data. By rich data, we mean a wide and diverse range of information collected over a relatively prolonged period of time in a persistent and systematic manner. Ideally, such data enable you to grasp the meanings associated with the actions of those you are studying and to understand the contexts in which those actions are embedded.*
>
> *(Lofland et al. 2006: 15)*

probability sampling The general term for samples selected in accordance with probability theory, typically involving some random-selection mechanism. Specific types of probability sampling include EPSEM, PPS, simple random sampling, and systematic sampling.

In other words, nonprobability sampling does have its uses, particularly in qualitative research projects. But researchers must take care to acknowledge the limitations of nonprobability sampling, especially regarding accurate and precise representations of populations. This point will become clearer as we discuss the logic and techniques of probability sampling.

The Logic and Techniques of Probability Sampling

Although appropriate to some research purposes, nonprobability-sampling methods cannot guarantee that the sample we observed is representative of the whole population. When researchers want precise, statistical descriptions of large populations—for example, the percentage of the population that is unemployed, that plans to vote for Candidate X, or that feels a rape victim should have the right to an abortion—they turn to **probability sampling**. All large-scale surveys use probability-sampling methods.

Although the application of probability sampling involves a somewhat sophisticated use of statistics, the basic logic of probability sampling is not difficult to understand. If all members of a population were identical in all respects—all demographic characteristics, attitudes, experiences, behaviors, and so on—there would be no need for careful sampling procedures. In this extreme case of perfect homogeneity, in fact, any single case would suffice as a sample to study characteristics of the whole population.

In fact, of course, the human beings who compose any real population are quite heterogeneous, varying in many ways. Figure 7-2 offers a simplified illustration of a heterogeneous population: The 100 members of this small population differ by gender and race. We'll use this hypothetical micropopulation to illustrate various aspects of probability sampling.

The fundamental idea behind probability sampling is this: In order to provide useful descriptions of the total population, a sample of individuals from a population must contain

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

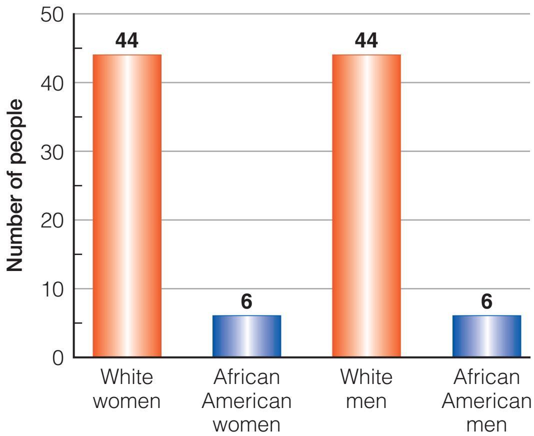


FIGURE 7-2

A Population of 100 Folks. Typically, sampling aims at reflecting the characteristics and dynamics of large populations. For the purpose of some simple illustrations, let's assume our total population has only 100 members.

essentially the same variations that exist in the population. This isn't as simple as it might seem, however. Let's take a minute to look at some of the ways researchers might go astray. Then, we'll see how probability sampling provides an efficient method for selecting a sample that should adequately reflect variations that exist in the population.

Conscious and Subconscious Sampling Bias

At first glance, it may look as though sampling is pretty straightforward. To select a sample of 100 university students, you might simply interview the first 100 students you find walking around campus. Although untrained researchers often use this kind of sampling method, it runs a high risk of introducing biases into the samples.

In connection with sampling, *bias* simply means that those selected are not typical or representative of the larger populations they've been chosen from. This kind of bias does not have to be intentional. In fact, it's virtually inevitable when you pick people by the seat of your pants.

Figure 7-3 illustrates what can happen when researchers simply select people who are convenient for study. Although women make up 50 percent of our micropopulation, the people closest to the researcher (in the lower right corner) happen to be 70 percent women, and although the population is 12 percent African American, none were selected for the sample.

Beyond the risks inherent in simply studying people who are convenient, other problems can arise. To begin, the researcher's personal leanings

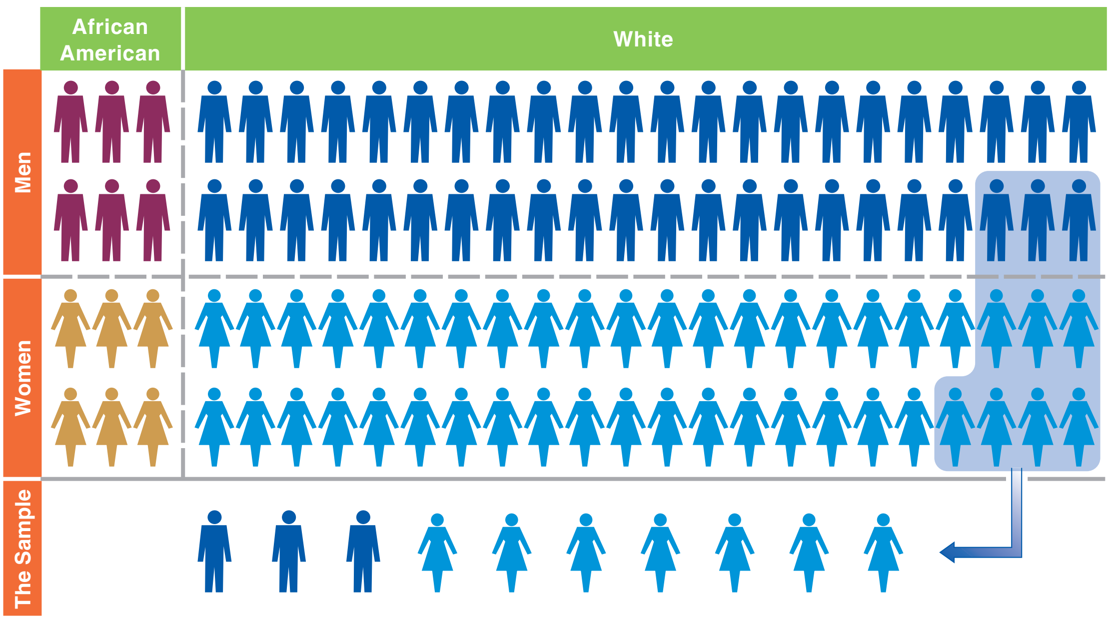


FIGURE 7-3

A Sample of Convenience: Easy, but Not Representative. Selecting and observing those people who are most readily at hand is the simplest method, perhaps, but it's unlikely to provide a sample that accurately reflects the total population.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

may affect the sample to the point where it does not truly represent the student population. Suppose you're a little intimidated by students who look particularly "cool," feeling that they might ridicule your research effort. You might consciously or subconsciously avoid interviewing such people. Or, you might feel that the attitudes of "super-straight-looking" students would be irrelevant to your research purposes, and so you avoid interviewing them.

Even if you sought to interview a "balanced" group of students, you wouldn't know the exact proportions of different types of students making up such a balance, and you wouldn't always be able to identify the different types just by watching them walk by.

Further, even if you made a conscientious effort to interview, say, every tenth student entering the university library, you could not be sure of a representative sample, because different types of students visit the library with different frequencies. Your sample would overrepresent students who visit the library more often than do others.

Similarly, the "public opinion" call-in polls—in which radio stations or newspapers ask people to call specified telephone numbers, text, or tweet to register their opinions—cannot be trusted to represent general populations. At the very least, not everyone in the population will even be aware of the poll. This problem also invalidates polls by magazines and newspapers who publish questionnaires for readers to complete and mail in. Even among those who are aware of such polls, not all will express an opinion, especially if doing so will cost them a stamp, an envelope, and their time. Similar considerations apply to polls taken over the Internet.

Ironically, the failure of such polls to represent all opinions equally was inadvertently acknowledged by Phillip Perinelli (1986), a staff manager of AT&T Communications' DIAL-IT 900 Service, which offers a call-in poll facility to organizations. Perinelli attempted to counter criticisms by saying, "The 50-cent charge assures that only interested parties respond and helps assure also that no individual 'stuffs' the ballot box." Social researchers cannot determine general public opinion while considering "only interested parties." This excludes those who don't care 50-cents' worth, as well as those who recognize that such polls are not valid. Both types of people may have opinions and may even vote on election day. Perinelli's assertion that the 50-cent charge will prevent ballot stuffing actually means that only those who can afford it will engage in ballot stuffing.

The possibilities for inadvertent sampling bias are endless and not always obvious. Fortunately, several techniques can help us avoid bias.

representativeness That quality of a sample of having the same distribution of characteristics as the population from which it was selected. By implication, descriptions and explanations derived from an analysis of the sample may be assumed to represent similar ones in the population. Representativeness is enhanced by probability sampling and provides for generalizability and the use of inferential statistics.

Representativeness and Probability of Selection

Although the term **representativeness** has no precise, scientific meaning, it carries a commonsense meaning that makes it useful here. For our purpose, a sample is representative of the population from which it is selected if the aggregate characteristics of the sample closely approximate those same aggregate characteristics in the population. If, for example, the population contains 50 percent women, then a sample must contain "close to" 50 percent women to be representative. Later, we'll discuss "how close" in detail. See "Applying Concepts in Everyday Life: Representative Sampling" for more on this.

Note that samples need not be representative in all respects; representativeness concerns only those characteristics that are relevant to the substantive interests of the study. However, you may not know in advance which characteristics are relevant.

A basic principle of probability sampling is that a sample will be representative of the population from which it is selected if all members of the population have an equal chance of being selected for the sample. (We'll see shortly that the size of the sample selected also affects the degree of

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Applying Concepts in Everyday Life

Representative Sampling

Representativeness applies to many areas of life, not just survey sampling. Consider quality control, for example. Imagine running a company that makes light bulbs. You want to be sure that they actually light up, but you can't test them all. You could, however, devise a method of selecting a sample of bulbs drawn from different times in the production day, on different machines, in different factories, and so forth.

Sometimes the concept of representative sampling serves as a protection against overgeneralization, discussed in Chapter 1. Suppose you go to a particular restaurant and don't like the food or service. You're ready to cross it off your list of dining possibilities, but then you think about it—perhaps you hit them on a bad night. Perhaps the chef had just discovered her boyfriend in bed with that "witch" from the Saturday wait staff and her mind wasn't on her cooking. Or perhaps the "witch" was serving your table and kept looking over her shoulder to see if anyone with a meat cleaver was bursting out of the kitchen. In short, your first experience might not have been representative.

representativeness.) Samples that have this quality are often labeled **EPSEM** samples (EPSEM stands for "equal probability of selection method"). Later we'll discuss variations of this principle, which forms the basis of probability sampling.

Moving beyond this basic principle, we must realize that samples—even carefully selected EPSEM samples—seldom, if ever, perfectly represent the populations from which they are drawn. Nevertheless, probability sampling offers two special advantages.

First, probability samples, although never perfectly representative, are typically more representative than other types of samples, because the biases previously discussed are avoided. In practice, a probability sample is more likely than a nonprobability sample to be representative of the population from which it is drawn.

Second, and more important, probability theory permits us to estimate the accuracy or representativeness of the sample. Conceivably, an uninformed researcher might, through wholly haphazard means, select a sample that nearly perfectly represents the larger population. The odds are against doing so, however, and we would be unable to estimate the likelihood that he or she has achieved representativeness. The probability sample, on the other hand, can provide an accurate estimate of success or failure. Shortly we'll see exactly how this estimate can be achieved.

I've said that probability sampling ensures that samples are representative of the population we wish to study. As we'll see in a moment, probability sampling rests on the use of a random-selection procedure. To develop this idea, though, we need to give more-precise meaning to two important terms: *element* and *population*.

An **element** is that unit about which information is collected and that provides the basis of analysis. Typically, in survey research, elements are people or certain types of people. However, other kinds of units can constitute the elements of social research: Families, social clubs, or corporations might be the elements of a study. In a given study, elements are often the same as units of analysis, though the former are used in sample selection and the latter in data analysis.

Up to now we've used the term *population* to mean the group or collection that we're interested in generalizing about. More formally, a **population** is the theoretically specified aggregation of study elements. Whereas the vague term *Americans* might be the target for a study, the delineation of the population would include the definition of the element "Americans" (for example, citizenship, residence) and the time referent for the study (Americans as of when?). Translating the abstract "adult New Yorkers" into a workable population would require a specification of the age defining *adult*

EPSEM (equal probability of selection method) A sample design in which each member of a population has the same chance of being selected for the sample.

element That unit of which a population is composed and that is selected for a sample. Elements are distinguished from units of analysis, which are used in data analysis.

population The theoretically specified aggregation of the elements in a study.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

and the boundaries of New York. Specifying "college student" would include a consideration of full- and part-time students, degree candidates and non–degree candidates, undergraduate and graduate students, and so forth.

A **study population** is that aggregation of elements from which the sample is actually selected. As a practical matter, researchers are seldom in a position to guarantee that every element meeting the theoretical definitions laid down actually has a chance of being selected in the sample. Even where lists of elements exist for sampling purposes, the lists are usually somewhat incomplete. Some students are always inadvertently omitted from student rosters. Some telephone subscribers have unlisted numbers.

Often, researchers decide to limit their study populations more severely than indicated in the preceding examples. National polling firms may limit their national samples to the 48 adjacent states, omitting Alaska and Hawaii for practical reasons. A researcher wishing to sample psychology professors may limit the study population to those in psychology departments, omitting those in other departments. Whenever the population under examination is altered in such fashion, you must make the revisions clear to your readers.

Random Selection

With these definitions in hand, we can define the ultimate purpose of sampling: to select a set of elements from a population in such a way that descriptions of those elements accurately portray the total population from which the elements are selected. Probability sampling enhances the likelihood of accomplishing this aim and also provides methods for estimating the degree of probable success.

Random selection is the key to this process. In **random selection**, each element has an equal chance of being selected independently of any other event in the selection process. Flipping a coin is the most frequently cited example: Provided that the coin is perfect (that is, not biased in terms of coming up heads or tails), the "selection" of a head or a tail is independent of previous selections of heads or tails. No matter how many heads turn up in a row, the chance that the next flip will produce "heads" is exactly 50–50. Rolling a perfect set of dice is another example.

Such images of random selection, though useful, seldom apply directly to sampling methods in social research. More typically, social researchers use tables of random numbers or computer programs that provide a random selection of sampling units. A **sampling unit** is that element or set of elements considered for selection at some stage of sampling. A little later, we'll see how computers are used to select random telephone numbers for interviewing, a technique called random-digit dialing.

There are two reasons for using random-selection methods. First, this procedure serves as a check on conscious or subconscious bias on the part of the researcher. The researcher who selects cases on an intuitive basis might very well select cases that would support his or her research expectations or hypotheses. Random selection erases this danger. Second, and more important, random selection offers access to the body of probability theory, which provides the basis for estimating the characteristics of the population as well as estimates of the precision of sample results. Now let's examine probability theory in greater detail.

Probability Theory, Sampling Distributions, and Estimates of Sampling Error

Probability theory is a branch of mathematics that provides the tools researchers need (1) to devise sampling techniques that produce representative samples and (2) to statistically analyze the results of their sampling. More formally, probability theory provides the basis for estimating the parameters of a population. A **parameter** is the summary description of a given variable in a population. The mean income of all families in a city is a

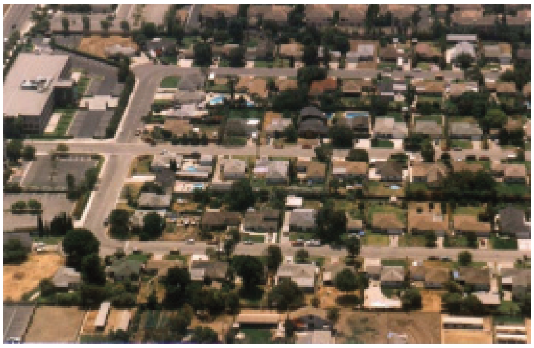
study population That aggregation of elements from which a sample is actually selected.

random selection A sampling method in which each element has an equal chance of being selected independently of any other event in the selection process.

sampling unit That element or set of elements considered for selection in some stage of sampling.

parameter The summary description of a given variable in a population.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.



Earl Babbie

How would researchers conduct a random sample of this neighborhood? What are the pitfalls they would need to avoid?

parameter; so is the age distribution of the city's population. When researchers generalize from a sample, they're using sample observations to estimate population parameters. Probability theory enables them to make these estimates and also to arrive at a judgment of how likely it is that the estimates will accurately represent the actual parameters in the population. So, for example, probability theory allows pollsters to infer from a sample of 2,000 voters how a population of 100 million voters is likely to vote—and to specify exactly what the probable margin of error in the estimates is.

Probability theory accomplishes these seemingly magical feats by way of the concept of sampling distributions. A single sample selected from a population will give an estimate of the population parameter. Other samples would give the same or slightly different estimates. Probability theory tells us about the distribution of estimates that would be produced by a large number of such samples.

The logic of sampling error can be applied to different kinds of measurements: mean income or mean age, for example. Measurements expressed as percentages, however, provide the simplest introduction to this general concept.

To see how this works, we'll look at two examples of sampling distributions, beginning with a simple example in which our population consists of just ten cases.

The Sampling Distribution of Ten Cases

Suppose that there are ten people in a group and that each has a certain amount of money in his or her pocket. To simplify, let's assume that one person has no money, another has one dollar, another has two dollars, and so forth up to the person with nine dollars. Figure 7-4 presents the population of ten people.

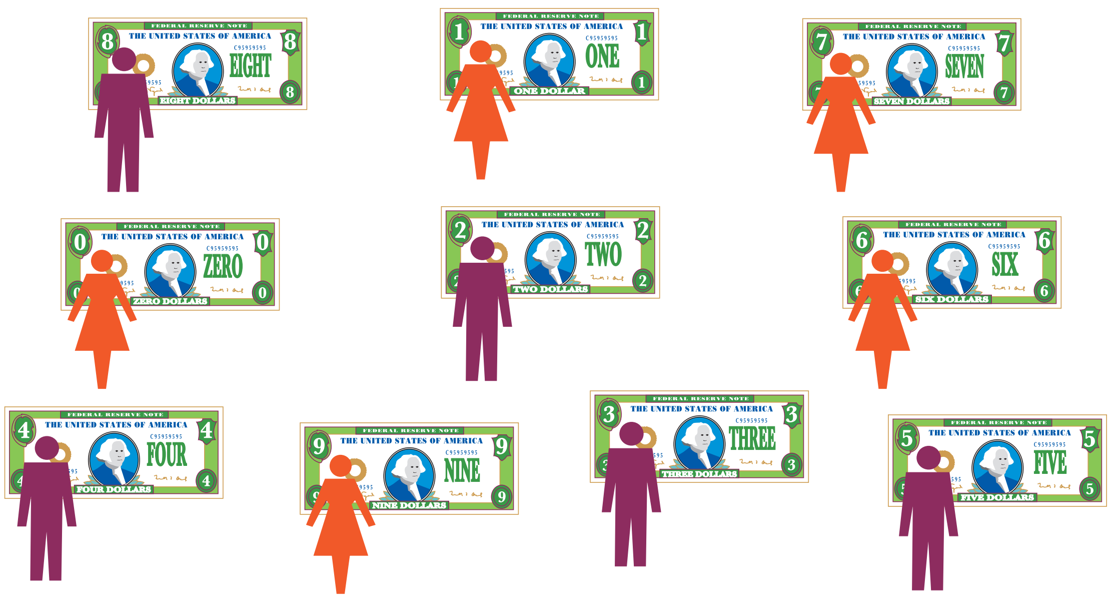


FIGURE 7-4

A Population of Ten People with $0 to $9. Let's imagine a population of only ten people with differing amounts of money in their pockets—ranging from $0 to $9.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Our task is to determine the average amount of money one person has: specifically, the mean number of dollars. If you simply add up the money shown in Figure 7-4, you'll find that the total is $45, so the mean is $4.50. Our purpose in the rest of this exercise is to estimate that mean without actually observing all ten individuals. We'll do that by selecting random samples from the population and using the means of those samples to estimate the mean of the whole population.

To start, suppose we were to select—at random—a sample of only one person from the ten. Our ten possible samples thus consist of the ten cases shown in Figure 7-4.

The ten dots shown on the graph in Figure 7-5 represent these ten samples. Because we're taking samples of only one, they also represent the "means" we would get as estimates of the population. The distribution of the dots on the graph is called the sampling distribution. Obviously, it wouldn't be a very good idea to select a sample of only one, because we'll very likely miss the true mean of $4.50 by quite a bit.

Now suppose we take a sample of two. As shown in Figure 7-6, increasing the sample size improves our estimations. There are now forty-five possible samples: [$0, $1], [$0, $2], . . . [$7, $8], [$8, $9]. Moreover, some of those samples produce the same means.

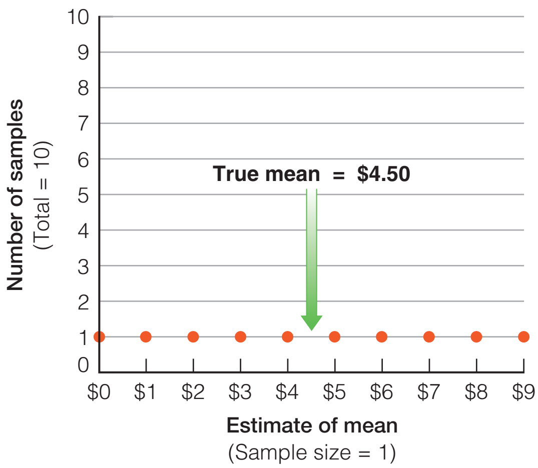


FIGURE 7-5

The Sampling Distribution of Samples of 1. In this simple example, the mean amount of money these people have is $4.50 ($45/10). If we picked ten different samples of one person each, our "estimates" of the mean would range across the board.

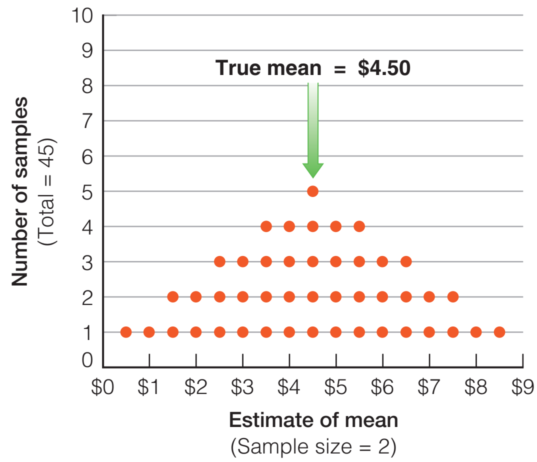


FIGURE 7-6

The Sampling Distribution of Samples of 2. After merely increasing our sample size to two, the possible samples provide somewhat better estimates of the mean. We couldn't get either $0 or $9, and the estimates are beginning to cluster around the true value of the mean: $4.50.

For example, [$0, $6], [$1, $5], and [$2, $4] all produce means of $3. In Figure 7-6, the three dots shown above the $3 mean represent those three samples.

Moreover, the forty-five samples are not evenly distributed, as they were when the sample size was only one. Rather, they cluster somewhat around the true value of $4.50. Only two possible samples deviate by as much as $4 from the true value ([$0, $1] and [$8, $9]), whereas five of the samples give the true estimate of $4.50; another eight samples miss the mark by only 50 cents (plus or minus).

Now suppose we select even larger samples. What do you suppose that will do to our estimates of the mean? Figure 7-7 presents the sampling distributions of samples of 3, 4, 5, and 6.

The progression of sampling distributions is clear. Every increase in sample size improves the distribution of estimates of the mean. The limiting case in this procedure, of course, is to select a sample of ten. There would be only one possible sample (everyone) and it would give us the true mean of $4.50. As we'll see shortly, this principle applies to actual sampling of meaningful populations. The larger the sample selected, the more accurate it is as an estimation of the population from which it was drawn.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

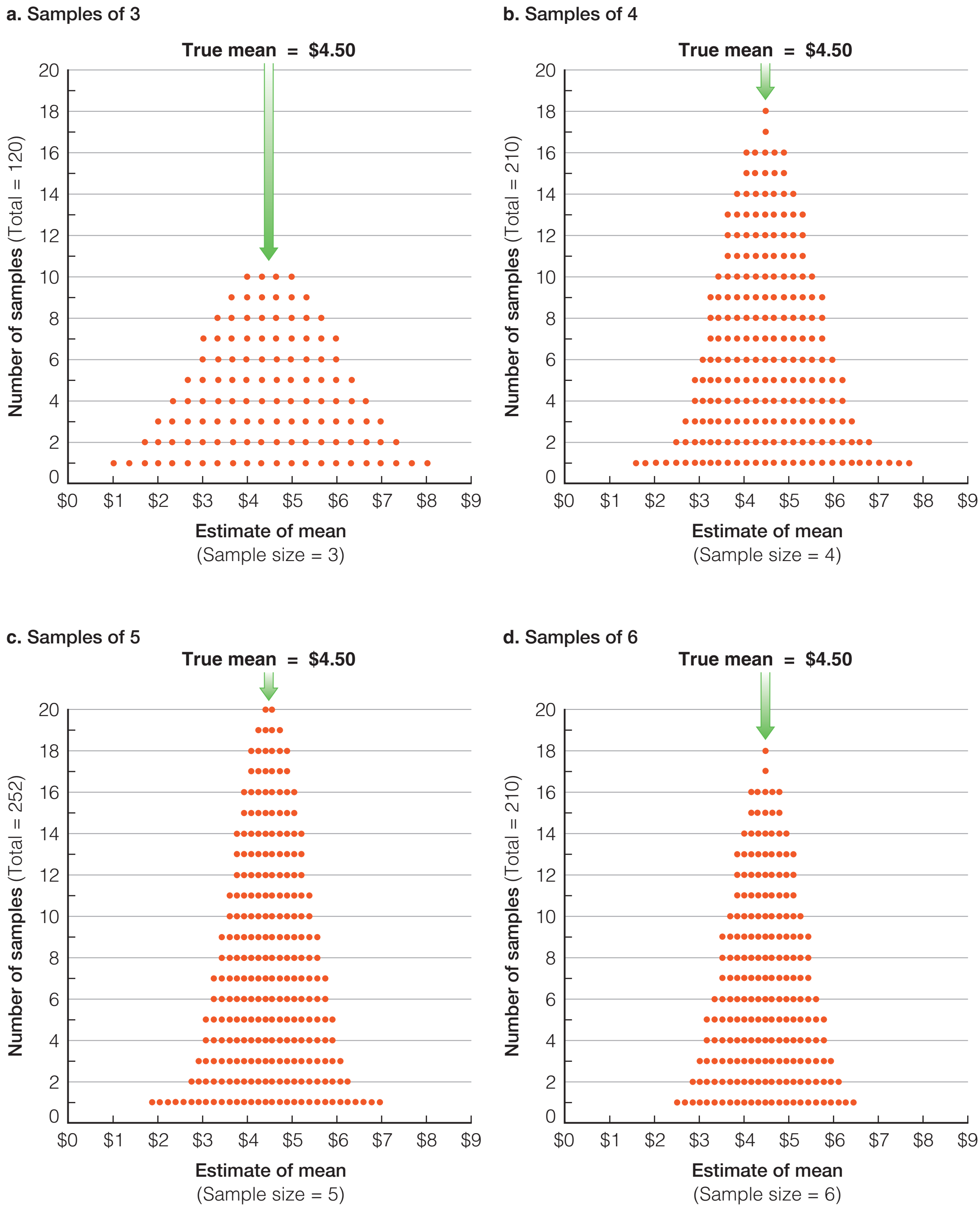


FIGURE 7-7

The Sampling Distributions of Samples of 3, 4, 5, and 6. As we increase the sample size, the possible samples cluster ever more tightly around the true value of the mean. The chance of extremely inaccurate estimates is reduced at the two ends of the distribution, and the percentage of the samples near the true value keeps increasing.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Sampling Distribution and Estimates of Sampling Error

Let's turn now to a more realistic sampling situation involving a much larger population and see how the notion of sampling distribution applies. Assume that we wish to study the student population of State University (SU) to determine the percentage of students who approve or disapprove of a student conduct code proposed by the administration. The study population will be the aggregation of, say, 20,000 students contained in a student roster: the sampling frame. The elements will be the individual students at SU. We'll select a random sample of, say, 100 students for the purposes of estimating the entire student body. The variable under consideration will be attitudes toward the code, a binomial variable comprising the attributes *approve* and *disapprove*. (The logic of probability sampling applies to the examination of other types of variables, such as *mean income*, but the computations are somewhat more complicated. Consequently, this introduction focuses on binomials.)

The horizontal axis of Figure 7-8 presents all possible values of this parameter in the population—from 0 percent to 100 percent approval. The midpoint of the axis—50 percent—represents half the students approving of the code and the other half disapproving.

To choose our sample, we give each student on the student roster a number and select 100 random numbers from a table of random numbers. Then we interview the 100 students whose numbers have been selected and ask whether they approve or disapprove of the student code. Suppose that this operation gives us forty-eight students who approve of the code and fifty-two who disapprove. This summary description of a variable in a sample is called a **statistic**. We present this statistic by placing a dot on the *x* axis at the point representing 48 percent.

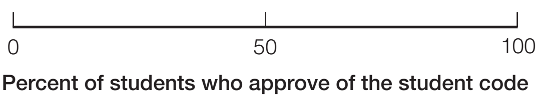


FIGURE 7-8

Range of Possible Sample Study Results. Shifting to a more realistic example, let's assume that we want to sample student attitudes concerning a proposed conduct code. Let's assume that 50 percent of the whole student body approves and 50 percent disapproves—though the researcher doesn't know that.

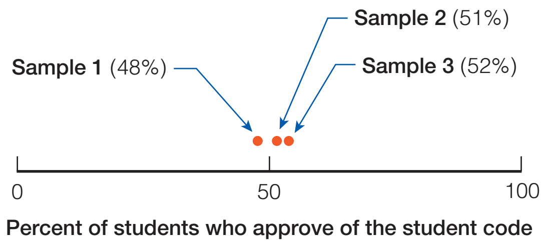


FIGURE 7-9

Results Produced by Three Hypothetical Studies. Assuming a large student body, let's suppose that we selected three different samples, each of substantial size. We would not necessarily expect those samples to perfectly reflect attitudes in the whole student body, but they should come reasonably close.

> **statistic** The summary description of a variable in a sample, used to estimate a population parameter.

Now let's suppose that we select another sample of 100 students in exactly the same fashion and measure their approval or disapproval of the student code. Perhaps fifty-one students in the second sample approve of the code. We place another dot in the appropriate place on the *x* axis. Repeating this process once more, we may discover that fifty-two students in the third sample approve of the code.

Figure 7-9 presents the three different sample statistics representing the percentages of students in each of the three random samples who approved of the student code. The basic rule of random sampling is that such samples, drawn from a population, give estimates of the parameter that exists in the total population. Each of the random samples, then, gives us an estimate of the percentage of students in the total student body who approve of the student code. Unhappily, however, we have selected three samples and now have three separate estimates.

To rescue ourselves from this problem, let's draw more and more samples of 100 students each, question each of the students in the samples concerning their approval or disapproval of the code, and plot the new sample statistics on our summary graph. In drawing many such samples, we discover that some of the new samples provide duplicate estimates, as in the illustration of ten cases. Figure 7-10 shows the sampling distribution of, say, hundreds of samples. This is often referred to as a *normal curve*.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

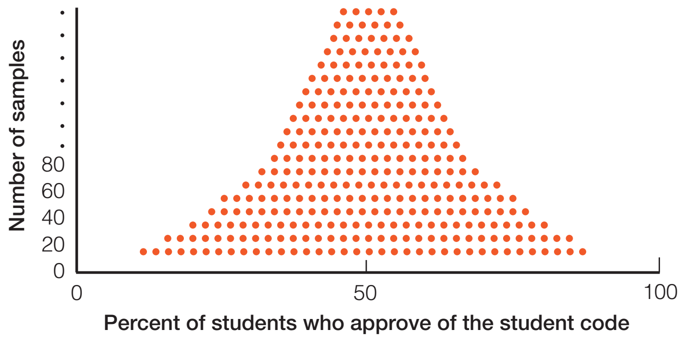


FIGURE 7-10

The Sampling Distribution. If we were to select a large number of good samples, we would expect them to cluster around the true value (50 percent), but given enough such samples, a few would fall far from the mark.

Note that by increasing the number of samples of students selected and interviewed, we've also increased the range of estimates provided by the sampling operation. In one sense we've increased our dilemma in attempting to guess the parameter in the population. Probability theory, however, provides certain important rules regarding the sampling distribution presented in Figure 7-10.

First, if many independent random samples are selected from a population, the sample statistics provided by those samples will be distributed around the population parameter in a known way. Thus, although Figure 7-10 shows a wide range of estimates, more of them fall near 50 percent than elsewhere in the graph. Probability theory tells us, then, that the true value is in the vicinity of 50 percent.

Second, probability theory gives us a formula for estimating how closely the sample statistics are clustered around the true value. To put it another way, probability theory enables us to estimate the **sampling error**—the degree of error to be expected for a given sample design. This formula contains three factors: the parameter (P), the sample size, and the standard error (a measure of sampling error):

$$s = \sqrt{\frac{P \times Q}{n}}$$

The P and Q in the formula equal the population parameters for the binomial: If 60 percent of the student body approves of the code and 40 percent disapproves, P and Q are 60 percent and 40 percent, respectively, or 0.6 and 0.4. The symbol n equals the number of cases in each sample, and s is the standard error.

Let's assume that the population parameter in the student example is 50 percent approving of the code and 50 percent disapproving. Recall that we've been selecting samples of 100 cases each. When these numbers are put into the formula, we find that the standard error equals 0.05, or 5 percent.

In probability theory, the *standard error* is a valuable piece of information because it indicates the extent to which the sample estimates will be distributed around the population parameter. (If you're familiar with the standard deviation in statistics, you may recognize that the standard error, in this case, is the standard deviation of the sampling distribution.) Specifically, probability theory indicates that certain proportions of the sample estimates will fall within specified increments—each equal to one standard error—from the population parameter. Approximately 34 percent (0.3413) of the sample estimates will fall within one standard error increment above the population parameter, and another 34 percent will fall within one standard error below the parameter. In our example, the standard error increment is 5 percent, so we know that 34 percent of our samples will give estimates of student approval between 50 percent (the parameter) and 55 percent (one standard error above);

sampling error The degree of error to be expected in probability sampling. The formula for determining sampling error contains three factors: the parameter, the sample size, and the standard error.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

another 34 percent of the samples will give estimates between 50 percent and 45 percent (one standard error below the parameter). Taken together, then, we know that roughly two-thirds (68 percent) of the samples will give estimates within 5 percentage points of the parameter.

Moreover, probability theory dictates that roughly 95 percent of the samples will fall within plus or minus two standard errors of the true value, and 99.9 percent of the samples will fall within plus or minus three standard errors. In our present example, then, we know that only one sample out of a thousand would give an estimate lower than 35 percent approval or higher than 65 percent.

The proportion of samples falling within one, two, or three standard errors of the parameter is constant for any random-sampling procedure such as the one just described, providing that a large number of samples are selected. The size of the standard error in any given case, however, is a function of the population parameter and the sample size. If we return to the formula for a moment, we note that the standard error will increase as a function of an increase in the quantity P times Q (PQ). Note further that this quantity reaches its maximum in the situation of an even split in the population. If $P = 0.5$, $PQ = 0.25$; if $P = 0.6$, $PQ = 0.24$; if $P = 0.8$, $PQ = 0.16$; if $P = 0.99$, $PQ = 0.0099$. By extension, if P is either 0.0 or 1.0 (either 0 percent or 100 percent approve of the student code), the standard error will be 0. If everyone in the population has the same attitude (no variation), then every sample will give exactly that estimate.

The standard error is also a function of the sample size—an inverse function. As the sample size increases, the standard error decreases. As the sample size increases, more and more samples will be clustered nearer to the true value. Another general guideline is evident in the formula: Because of the square-root formula, the standard error is reduced by half if the sample size is quadrupled.

confidence level The estimated probability that a population parameter lies within a given confidence interval. Thus, we might be 95 percent confident that between 35 and 45 percent of all voters favor Candidate A.

confidence interval The range of values within which a population parameter is estimated to lie.

In our present example, samples of 100 produce a standard error of 5 percent; to reduce the standard error to 2.5 percent, we must increase the sample size to 400.

All of this information is provided by established probability theory in reference to the selection of large numbers of random samples. (If you've taken a statistics course, you may know this as the central tendency theorem.) If the population parameter is known and many random samples are selected, we can predict how many of the sampling estimates will fall within specified intervals from the parameter.

Recognize that this discussion illustrates only the logic of probability sampling; it does not describe the way research is actually conducted. Usually, we don't know the parameter: The very reason we conduct a sample survey is to estimate that value. Moreover, we don't actually select large numbers of samples: We select only one sample. Nevertheless, the preceding discussion of probability theory provides the basis for inferences about the typical social research situation. Knowing what it would be like to select thousands of samples allows us to make assumptions about the one sample we do select and study.

Confidence Levels and Confidence Intervals

Whereas probability theory specifies that 68 percent of that fictitious large number of samples would produce estimates falling within one standard error of the parameter, we can turn the logic around and infer that any single random sample has a 68 percent chance of falling within that range. This observation leads us to the two key components of sampling-error estimates: **confidence level** and **confidence interval**. We express the precision of our sample statistics in terms of a level of confidence that the statistics fall within a specified interval from the parameter. For example, we may say we are 95 percent confident that our sample statistics (for example, 50 percent favor the new student code) are within plus or minus 10 percentage points of the population parameter. As the confidence interval is expanded for a given statistic, our confidence increases. For example, we may say that we are 99.9 percent confident that our statistic falls within three standard errors of the true value. (Now perhaps you can appreciate the humorous

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

quip of unknown origin: Statistics means never having to say you are certain.)

Although we may be confident (at some level) of being within a certain range of the parameter, we've already noted that we seldom know what the parameter is. To resolve this problem, we substitute our sample estimate for the parameter in the formula; that is, lacking the true value, we substitute the best available guess.

The result of these inferences and estimations is that we can estimate a population parameter and also the expected degree of error on the basis of one sample drawn from a population. Beginning with the question "What percentage of the student body approves of the student code?" you could select a random sample of 100 students and interview them. You might then report that your best estimate is that 50 percent of the student body approves of the code and that you are 95 percent confident that between 40 and 60 percent (plus or minus two standard errors) approve. The range from 40 to 60 percent is the confidence interval. (At the 68 percent confidence level, the confidence interval would be 45 to 55 percent.)

The logic of confidence levels and confidence intervals also provides the basis for determining the appropriate sample size for a study. Once you've decided on the degree of sampling error you can tolerate, you'll be able to calculate the number of cases needed in your sample. Thus, for example, if you want to be 95 percent confident that your study findings are accurate within plus or minus 5 percentage points of the population parameters, you should select a sample of at least 400. (Appendix E is a convenient guide in this regard.)

This, then, is the basic logic of probability sampling. Random selection permits the researcher to link findings from a sample to the body of probability theory so as to estimate the accuracy of those findings. All statements of accuracy in sampling must specify both a confidence level and a confidence interval. The researcher must report that he or she is *x* percent confident that the population parameter lies between two specific values. In this example, I've demonstrated the logic of sampling error using a variable analyzed in percentages. Although different statistical procedures would be required to calculate the standard error for a mean, for example, the overall logic is the same.

Notice that nowhere in this discussion of sample size and accuracy of estimates did we consider the size of the population being studied. This is because the population size is almost always irrelevant. A sample of 2,000 respondents drawn properly to represent Vermont voters will be no more accurate than a sample of 2,000 drawn properly to represent all voters in the United States, even though the Vermont sample would be a substantially larger proportion of that small state's voters than would the same number chosen to represent the nation's voters. The reason for this counterintuitive fact is that the equations for calculating sampling error all assume that the populations being sampled are infinitely large, so every sample would equal 0 percent of the whole.

Of course, this is not literally true in practice. A sample of 2,000 represents 0.63 percent of the Vermonters who voted for president in the 2016 election, and a sample of 2,000 U.S. voters represents 0.0014 percent of the national electorate. Nonetheless, both of these proportions are small enough to approach the ideal of a sample taken from an infinitely large population. Further, that proportion remains irrelevant unless a sample represents, say, 5 percent or more of the population it's drawn from. In those rare cases of large proportions being selected, a "finite population correction" can be calculated to adjust the confidence intervals.

The following formula calculates the proportion to be multiplied against the calculated error:

$$\text{finite population correction} = \sqrt{\frac{N - n}{N - 1}}$$

In the formula, N is the population size and n is the size of the sample. Notice that in the extreme case where you studied the whole population (hence, $N - n$), the formula would yield zero as the finite population correction. Multiplying zero times the sampling error calculated by the earlier formula would give a final sampling error of zero, which would, of course, be precisely the case because you wouldn't have sampled at all.

Lest you weary of the statistical nature of this discussion, it's useful to realize what an amazing thing we've been examining. There is remarkable order within what might seem random and chaotic. One of the researchers to whom we owe this observation is Sir Francis Galton (1822–1911):

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

> *Order in Apparent Chaos. I know of scarcely anything so apt to impress the imagination as the wonderful form of cosmic order expressed by the "Law of Frequency of Error." The law would have been personified by the Greeks and deified, if they had known of it. It reigns with serenity and in complete self-effacement amidst the wildest confusion. The huger the mob, and the greater the apparent anarchy, the more perfect is its sway. It is the supreme law of Unreason.*
>
> *(Galton 1889: 66)*

Two cautions are in order before we conclude this discussion of the basic logic of probability sampling. First, the survey uses of probability theory as discussed here are technically not wholly justified. The theory of sampling distribution makes assumptions that almost never apply in survey conditions. The exact proportion of samples contained within specified increments of standard errors, for example, mathematically assumes an infinitely large population, an infinite number of samples, and sampling with replacement—that is, every sampling unit selected is "thrown back into the pot" and could be selected again. Second, our discussion has greatly oversimplified the inferential jump from the distribution of several samples to the probable characteristics of one sample.

I offer these cautions to provide perspective on the uses of probability theory in sampling. Social researchers often appear to overestimate the precision of estimates produced by the use of probability theory. As I'll mention elsewhere in this chapter and throughout the book, variations in sampling techniques and nonsampling factors may further reduce the legitimacy of such estimates. For example, those selected in a sample who fail or refuse to participate further detract from the representativeness of the sample.

Over the years, the public has grown accustomed to having the media reports ranges of sampling error to accompany political poll results, for example. However, there is still room for improvement in this regard. During the 2015–2016 presidential campaigns, there were as many as seventeen Republican candidates at one point. This was a problem for the logistics of televised debates. The general solution was to limit the debates to the most popular candidates, leaving out those who had little or no public support. But how was the popularity decision to be made?

Each of the debate lineups was selected somewhat differently, but the basic logic was to determine the target number, say ten, and decide on a set of national polls that would be conducted prior to the debate. Each candidate would receive a score equal to the average popularity that candidate received in the set of specified polls. As reasonable as this might seem, the difference in popularity scores of candidate 10 and candidate 11 was always well within the margin of sampling error, sometimes a fraction of a percentage point. It was a meaningless statistical difference but a profound political difference.

Lest this seem a partisan criticism of the Republican Party, the Democrats used precisely the same hair-splitting technique in the run-up to their 2020 Presidential primary.

Though they can be misused, the calculations discussed in this section can be extremely valuable to you in understanding and evaluating your data. Although the calculations do not provide estimates as precise as some researchers might assume, they can be quite valid for practical purposes. They are unquestionably more valid than less rigorously derived estimates based on less rigorous sampling methods. Most important, being familiar with the basic logic underlying the calculations can help you react sensibly both to your own data and to those reported by others.

Populations and Sampling Frames

The preceding section introduced the theoretical model for social research sampling. Although as students, research consumers, and researchers we need to understand that theory, appreciating the less-than-perfect conditions that exist in the field is no less important. In this section we'll look at one aspect of field conditions that requires a compromise with idealized theoretical conditions and assumptions: the congruence of or disparity between populations of sampling frames.

Simply put, a **sampling frame** is the list or quasi-list of elements from which a probability

sampling frame The list or quasi-list of units composing a population from which a sample is selected. If the sample is to be representative of the population, it is essential that the sampling frame include all (or nearly all) members of the population.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

sample is selected. If a sample of students is selected from a student roster, the roster is the sampling frame. If the primary sampling unit for a complex population sample is the census block, the list of census blocks composes the sampling frame—in the form of a printed booklet or, better, some digital format permitting computer manipulation. Here are some reports of sampling frames appearing in research journals. In each example I've italicized the actual sampling frames:

> *The data for this research were obtained from a random sample of* parents of children in the third grade in public and parochial schools in Yakima County, Washington.
>
> *(Petersen and Maynard 1981: 92)*

> *The sample at Time 1 consisted of 160 names drawn randomly from the* telephone directory of Lubbock, Texas.
>
> *(Tan 1980: 242)*

> *The data reported in this paper . . . were gathered from a probability sample of* adults aged 18 and over residing in households in the 48 contiguous United States. *Personal interviews with 1,914 respondents were conducted by the Survey Research Center of the University of Michigan during the fall of 1975.*
>
> *(Jackman and Senter 1980: 345; emphasis mine)*

Properly drawn samples provide information appropriate for describing the population of elements composing the sampling frame—nothing more. I emphasize this point in view of the all-too-common tendency for researchers to select samples from a given sampling frame and then make assertions about a population similar to, but not identical to, the population defined by the sampling frame.

For example, take a look at this report, which discusses the drugs most frequently prescribed by U.S. physicians:

> *Information on prescription drug sales is not easy to obtain. But Rinaldo V. DeNuzzo, a professor of pharmacy at the Albany College of Pharmacy, Union University, Albany, NY, has been tracking prescription drug sales for 25 years by polling nearby drugstores. He publishes the results in an industry trade magazine,* MM&M.
>
> *DeNuzzo's latest survey, covering 1980, is based on reports from 66 pharmacies in 48 communities in New York and New Jersey. Unless there is something peculiar about that part of the country, his findings can be taken as representative of what happens across the country.*
>
> *(Moskowitz 1981: 33)*

What is striking in the excerpt is the casual comment about whether there is anything peculiar about New York and New Jersey. There is. The lifestyle in these two states hardly typifies the lifestyles in the other 48. We cannot assume that residents in these large, urbanized, Eastern Seaboard states necessarily have the same prescription-drug–use patterns that residents of Mississippi, Nebraska, or Vermont have.

Does the survey even represent prescription patterns in New York and New Jersey? To determine that, we would have to know something about the way the 48 communities and the 66 pharmacies were selected. We should be wary in this regard, in view of the reference to "polling nearby drugstores." As we'll see, there are several methods for selecting samples that ensure representativeness, and unless they're used, we shouldn't generalize from the study findings.

A sampling frame, then, must be consonant with the population we wish to study. In the simplest sample design, the sampling frame is a list of the elements composing the study population. In practice, though, existing sampling frames often define the study population rather than the other way around. That is, we often begin with a population in mind for our study; then we search for possible sampling frames. Having examined and evaluated the frames available for our use, we decide which frame presents a study population most appropriate to our needs.

Studies of organizations are often the simplest from a sampling standpoint because organizations typically have membership lists. In such cases, the list of members constitutes an excellent sampling frame. If a random sample is selected from a membership list, the data collected from that sample may be taken as representative of all members—if all members are included in the list.

Populations that can be sampled from good organizational lists include elementary school, high school, and university students and faculty; church members; factory workers; fraternity or sorority members; members of social, service, or political clubs; and members of professional associations.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

The preceding comments apply primarily to local organizations. Often, statewide or national organizations do not have a single membership list. There is, for example, no single list of Episcopalian church members. However, a slightly more complex sample design could take advantage of local church membership lists by first sampling churches and then subsampling the membership lists of the churches selected. (More about that later.)

Other lists that may be available contain the names of automobile owners, welfare recipients, taxpayers, business-permit holders, licensed professionals, and so forth. Although it may be difficult to gain access to some of these lists, they provide excellent sampling frames for specialized research purposes.

Certain lists of individuals may be especially relevant to the research needs of a particular study. For example, government agencies maintain lists of registered voters, and some political pollsters use those lists to do registration-based sampling. (In some cases, however, such files are not up to date; further, a person who is registered to vote may not actually vote in a given election.)

The sampling elements in a study need not be individual persons. Lists of other types of elements also exist: universities, corporations, cities, academic journals, newspapers, unions, political clubs, professional associations, and so forth.

Telephone directories were once used for "quick-and-dirty" public opinion polls. It's undeniable that they're easy and inexpensive to use—no doubt the reason for their popularity. And, if you want to make assertions about telephone subscribers, a directory is a fairly good sampling frame. (Realize, of course, that a given directory will include neither new subscribers nor those who have requested unlisted numbers, or cell phone numbers. Sampling is further complicated by the directories' inclusion of nonresidential listings.)

The earliest telephone surveys had a rather bad reputation among professional researchers. Telephone surveys are limited by definition to people who have telephones. Years ago, this method produced a substantial social-class bias by excluding poor people from the surveys. This was vividly demonstrated by the *Literary Digest* fiasco of 1936. Recall that, even though voters were contacted by mail, the sample was partially selected from telephone subscribers, who were hardly typical in a nation struggling with the Great Depression. By 2009, however, 95.7 percent of all households had telephones, so the earlier form of class bias has virtually disappeared (U.S. Bureau of the Census 2012: 712, Table 1132).

A related sampling problem involved unlisted numbers. A survey sample selected from the pages of a local telephone directory would totally omit all those people—typically richer—who requested that their numbers not be published. This potential bias was erased through a technique that has advanced telephone sampling substantially: random-digit dialing (RDD).

Imagine selecting a set of seven-digit telephone numbers at random. Even people whose numbers were unlisted would have the same chance of selection as would those in the directory. However, if you simply dialed randomly selected numbers, a high proportion of those would turn out to be "not in service," government offices, commercial enterprises, and so forth. Fortunately, it's possible to obtain ranges of numbers that are mostly active residential numbers. Selecting a set of those numbers at random will provide a representative sample of residential households. As a consequence, RDD has become a standard procedure in telephone surveys.

The growth in popularity of cell phones has created a new source of concern for survey researchers. The Telephone Consumer Protection Act of 1991 put limitations on telephone solicitations and, because calls to a cell phone may incur an expense to the target of the call (depending on their service plan), the Act made it illegal for automatic dialing systems (e.g., the robocalls alerting you to a special sale on widgets) to call cell phones (FCC 2015). But where does this leave survey researchers, who aren't selling anything? While efforts are underway to officially exempt research projects from that ruling, AAPOR (2010) advises members that:

> *To ensure compliance with this federal law, in the absence of express prior consent from a sampled cell phone respondent, telephone research call centers should have their interviewers* manually dial cell phone numbers *(i.e., where a human being physically touches the numerals on the telephone to dial the number).*

Those who use cell phones exclusively, moreover, tend to be younger than the general population. In 2004, they were more likely to vote for John Kerry than older voters were.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

In 2008, they were more likely than the average voter to support Barack Obama. In a study of this matter, Scott Keeter and Kylie McGeeney (2015) found that a distinct bias by age and the variables closely related to it (such as marital status) distinguished those who were reachable only by cell phone and those reachable by landline.

> *For example, young adults, Hispanics, renters and the poor (as defined by the U.S. Census Bureau's poverty thresholds) are all far more likely to be cell-only. To the extent that cell-only households are underrepresented in our samples, these groups are also underrepresented.*
>
> *(Keeter and McGeeney, 2015)*

At the 2008 meeting of the American Association for Public Opinion Research (AAPOR), several research papers examined the implications of cell-phone popularity. Overall, most of the researchers found that ignoring people who use only cell phones did not seriously bias survey results, in most cases; this is because these people represented a relatively small portion of all telephone customers. That situation has changed quickly and substantially, as reported by the National Center for Health Statistics (2017):

> *The second 6 months of 2016 was the first time that a majority of American homes had only wireless telephones. Preliminary results from the July–December 2016 National Health Interview Survey (NHIS) indicate that 50.8% of American homes did not have a landline telephone but did have at least one wireless telephone. . .*

In part, researchers have sought to address the dramatic increase in cell phones by augmenting RDD sampling with address-based sampling (ABS), based on U.S. Postal Service lists of residential addresses. It will be possible for researchers to match some of the residential addresses with phone numbers by using directories that contain both pieces of information. The nonmatched addresses will be mailed a request for a phone number that researchers can then call. The mailing will also include a 1-800 number that potential survey respondents can call if they have no telephone number of their own. If two sampling frames are employed, however, it is important to either (1) rule out duplicate residences before sampling or (2) identify respondents who have both cell phones and landlines so that their responses can be weighted half as much as those with only one chance of being selected into the sample. The preferred method is still under study and debate (Boyle, Lewis, and Tefft 2010).

© arek_malang/Shutterstock.com

Cell phones have complicated survey sampling.

In an updated report from the Pew Center (Christian et al. 2010), special attention was paid to differences in the estimation of opinions and behaviors using landline-only samples and dual-frame samples, including cell-phone-only households.

> *The items selected include nearly all of the key indicators regularly tracked by our two centers (e.g., presidential approval, party affiliation, internet use, broadband adoption, sending and receiving text messages on a cell phone), as well as a sampling of other important measures that were timely or are asked intermittently (e.g., agreement with the Tea Party, approval of health care legislation, use of cell phones to play music).*
>
> *(2010: 2)*

Overall, the Pew researchers found relatively little bias due to sample design.

> *Despite the growth in cell-only households, the magnitude of possible non-coverage bias remains relatively small for the majority of measures tested. Of 72 questions examined, 43 of them show differences of 0, 1 or 2 percentage points between the landline and dual frame weighted samples.*
>
> *(2010: 2)*

This is an issue that will be followed closely by survey researchers in the years to come, as cell phones, presumably, will become ever-more dominant.

If cell phones have created a problem for sampling frames and sampling, that challenge pales in comparison with the new interest in online surveys. As we shall see in Chapter 9, there are substantial advantages to conducting surveys via the Web, but obtaining responses

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

from a sample representing the population of interest (e.g., adult population, voters) is especially tricky. At the most basic, not everyone participates on the Internet, comparable to the problem early in telephone polls, when not everyone had a telephone. Beyond that basic problem, those who are active on the Web participate to differing degrees and visit different websites.

Though a variety of methods are being tested at present, there is no clear solution to the problem of representativeness. As mentioned earlier in this chapter, we may see a rebirth of quota-sampling techniques as a way of making online samples represent the larger populations we are interested in studying. Another way of framing this issue is in terms of sampling error, as discussed earlier.

The calculation of sampling error, often called the *margin of error* (MOE), has a solid statistical grounding, providing it satisfies two criteria: (1) a probability sample is selected from a complete listing of the population of interest, and (2) everyone selected participates in the study. Now the bad news: rarely is either of these criteria fully satisfied. In our discussion of sampling frames, we saw that it is often difficult to obtain a complete listing or quasi-listing of the entire population under study. In the case of telephone surveys, we have seen the difficulty in accommodating the proliferation of cell phones. And the second criterion has been steadily undercut by falling response rates. Thus, survey researchers struggle to weigh their results to arrive at representative estimates of the population.

This problem is even greater in the case of nonprobability, "opt-in" polls on the Web. Potential respondents are invited to visit a website to participate in a poll. Some do, some don't, but there is no way of calculating conventional probability statistics regarding sampling error, because no probability sample was selected. In response to this dilemma, you may find reports of "credibility intervals," which resemble margins of sampling error in form (e.g., 95 percent confident of a 3-percentage-point error). Such estimates are based on the researchers choosing models that they believe distinguish opt-in respondents and the larger population of interest. This is a relatively new technique and somewhat controversial among researchers. The American Association for Public Opinion Research suggests caution in using or accepting credibility intervals.

Review of Populations and Sampling Frames

Because social research literature gives surprisingly little attention to the issues of populations and sampling frames, I've devoted special consideration to them here by providing a summary of the main guidelines to remember:

1. Findings based on a sample can be taken as representing only the aggregation of elements that compose the sampling frame.
2. Often, sampling frames do not truly include all the elements their names might imply. Omissions are almost inevitable. Thus, a first concern of the researcher must be to assess the extent of the omissions and to correct them if possible. (Of course, the researcher may feel that he or she can safely ignore a small number of omissions that cannot easily be corrected.)
3. To be generalized even to the population composing the sampling frame, all elements must have equal representation in the frame. Typically, each element should appear only once. Elements that appear more than once will have a greater probability of selection, and the sample will, overall, overrepresent those elements.

Other, more practical matters relating to populations and sampling frames will be treated elsewhere in this book. For example, the form of the sampling frame—such as a list in a publication, 3-by-5 card file, hard-drive file, or other digital storage method—can affect how easy it is to use. And ease of use may often take priority over scientific considerations: An "easier" list may be chosen over a "harder" one, even though the latter is more appropriate to the target population. Every researcher should carefully weigh the relative advantages and disadvantages of such alternatives.

Types of Sampling Designs

Up to this point, we've focused on simple random sampling. Indeed, the body of statistics typically used by social researchers assumes such a sample. As you'll see shortly, however, you have several options in choosing your sampling method, and you'll seldom if ever choose simple random sampling. There are two reasons for this. First, with all but the simplest sampling frame, simple random sampling is not feasible. Second, and

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

How to Do It

Using a Table of Random Numbers

In social research, it's often appropriate to select a set of random numbers from a table such as the one in Appendix B. Here's how to do that.

Suppose you want to select a simple random sample of 400 people (or other units) out of a population totaling 9,300.

1. To begin, number the members of the population: in this case, from 1 to 9,300. Now the problem is to select 400 random numbers. Once you've done that, your sample will consist of the people having the numbers you've selected. (*Note:* It's not essential to actually number them, as long as you're sure of the total. If you have them in a list, for example, you can always count through the list after you've selected the numbers.)
2. The next step is to determine the number of digits you'll need in the random numbers you select. In our example, there are 9,300 members of the population, so you'll need four-digit numbers to give everyone a chance of selection. (If there were 11,825 members of the population, you'd need to select five-digit numbers.) Thus, we want to select 400 random numbers in the range from 0001 to 9300.
3. On the first page of Appendix B, note that there are several rows and columns of five-digit numbers and that there are two pages. The table represents a series of random numbers in the range from 00001 to 99999. To use the table for your hypothetical sample, you have to answer these questions:

   a. How will you create four-digit numbers out of five-digit numbers?

   b. What pattern will you follow while moving through the table to select your numbers?

   c. Where will you start?

   Each of these questions has several satisfactory answers. The key is to create a plan and follow it. Here's an example.
4. To create four-digit numbers from five-digit numbers, let's agree to select five-digit numbers from the table but consider only the leftmost four digits in each case. If we picked the first number on the first page—51426—we would only consider the 5142. (We could agree to take the digits farthest to the right, 1426.) The key is to make a plan and stick with it. For convenience, let's use the leftmost four digits.
5. We can also choose to progress through the tables in any way we want: down the columns, up them, left to right or right to left, or diagonally. Again, any of these plans will work fine as long as we stick to it. For convenience, let's agree to move down the columns. When we get to the bottom of one column, we'll go to the top of the next.
6. Now, where do we start? You can close your eyes and stick a pencil into the table and start wherever the pencil point lands. (I know it doesn't sound scientific, but it works.) Or, if you're afraid you'll hurt the book or miss it altogether, close your eyes and make up a column number and a row number. ("I'll pick the number in the fifth row of column 2.") Start with that number.
7. Let's suppose we decide to start with the sixth number in column 2. On the first page of Appendix B, you'll see that the starting number is 09599. We have selected 0959 as our first random number, and we have 399 more to go. Moving down the second column, we select 3333, 1342, 3695, 4484, 7074, 3158, 9435—that's a problem, as there are only 9,300 people in the population. The solution is simple: ignore that number. If you happen to get the same number twice, ignore it the second time. Technically, this is called "random selection without replacement." After skipping 9435, we proceed to 6661, 4592, and so forth. When we get to the bottom of column 2, move to the top of column 3.
8. That's it. You keep up the procedure until you've selected 400 random numbers. Returning to your list, your sample consists of person number 0959, person number 3333, person number 1342, and so forth.

probably surprisingly, simple random sampling may not be the most accurate method available. Let's turn now to a discussion of simple random sampling and the other available options.

Simple Random Sampling

As noted, **simple random sampling** is the basic sampling method assumed in the statistical computations of social research. Because the mathematics of random sampling are especially complex, we'll detour around them in favor of describing the ways of employing this method in the field.

Once a sampling frame has been properly established, to use simple random sampling the researcher assigns a single number to each element in the list, not skipping any number in the process. A table of random numbers (Appendix B) is then used to select elements for the sample. "How to Do It: Using a Table of Random Numbers" explains its use.

simple random sampling A type of probability sampling in which the units composing a population are assigned numbers. A set of random numbers is then generated, and the units having those numbers are included in the sample.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

If your sampling frame is in a machine-readable form, such as a hard-drive file or flash drive, a simple random sample can be selected automatically by computer. (In effect, the computer program numbers the elements in the sampling frame, generates its own series of random numbers, and prints out the list of elements selected.)

Figure 7-11 offers an illustration of simple random sampling. Note that the members of our hypothetical micropopulation have been numbered from 1 to 100. Moving to the small table of random numbers provided, we decide to use the last two digits of the third column and to begin with the third number from the top. This yields person number 12 as the first one selected for the sample. Number 97 is next, and so forth. (Person 100 would have been selected if "00" had come up in the list.)

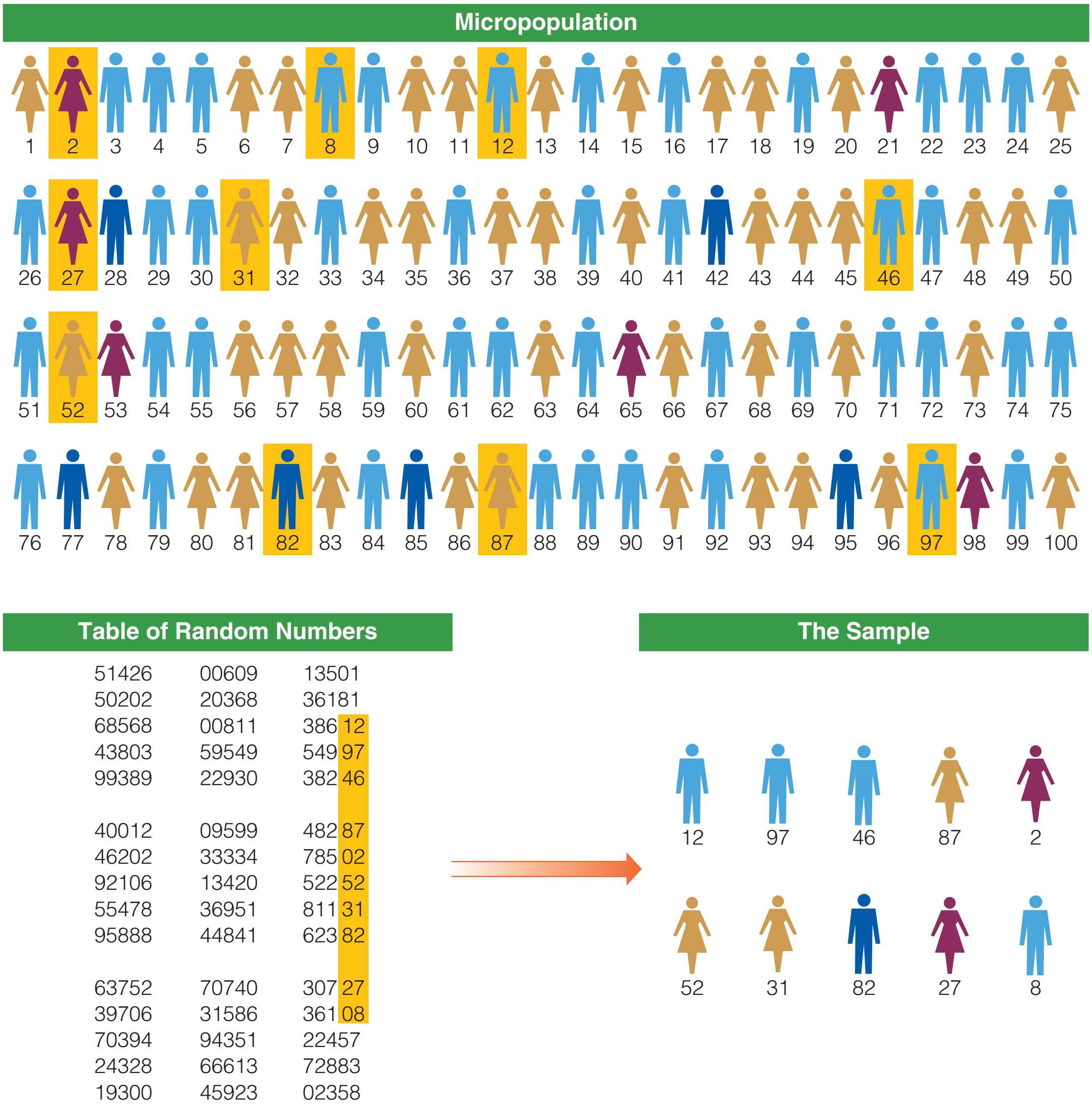


FIGURE 7-11

A Simple Random Sample. Having numbered everyone in the population, we can use a table of random numbers to select a representative sample from the overall population. Anyone whose number is chosen from the table is in the sample.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Systematic Sampling

Simple random sampling is seldom used in practice. As you'll see, it's not usually the most efficient method, and it can be laborious if done manually. Typically, simple random sampling requires a list of elements. When such a list is available, researchers usually employ systematic sampling instead.

In **systematic sampling**, every *k*th element in the total list is chosen (systematically) for inclusion in the sample. If the list contained 10,000 elements and you wanted a sample of 1,000, you would select every tenth element for your sample. To ensure against any possible human bias in using this method, you should select the first element at random. Thus, in the preceding example, you would begin by selecting a random number between one and ten. The element having that number is included in the sample, as is every tenth element following it. This method is technically referred to as a *systematic sample with a random start*. Two terms are frequently used in connection with systematic sampling. The **sampling interval** is the standard distance between elements selected in the sample: ten in the preceding sample. The **sampling ratio** is the proportion of elements in the population that are selected: 1/10 in the example.

$$\text{sampling interval} = \frac{\text{population size}}{\text{sample size}}$$

$$\text{sampling ratio} = \frac{\text{sample size}}{\text{population size}}$$

In practice, systematic sampling is virtually identical to simple random sampling. If the list of elements is indeed randomized before sampling, one might argue that a systematic sample drawn from that list is in fact a simple random sample. By now, debates over the relative merits of simple random sampling and systematic sampling have been resolved largely in favor of the latter, simpler method. Empirically, the results are virtually identical. And, as you'll see in a later section, systematic sampling, in some instances, is slightly more accurate than simple random sampling.

There is one danger involved in systematic sampling. The arrangement of elements in the list can make systematic sampling unwise. Such an arrangement is usually called *periodicity*. If the list of elements is arranged in a cyclical pattern that coincides with the sampling interval, a grossly biased sample can be drawn. Here are two examples that illustrate this danger.

In a classic study of soldiers during World War II, the researchers selected a systematic sample from unit rosters. Every tenth soldier on the roster was selected for the study. The rosters, however, were arranged in a table of organizations: sergeants first, then corporals and privates, squad by squad. Each squad had ten members. As a result, every tenth person on the roster was a squad sergeant. The systematic sample selected contained only sergeants. It could, of course, have been the case that no sergeants were selected for much the same reason.

As another example, suppose we select a sample of apartments in an apartment building. If the sample is drawn from a list of apartments arranged in numerical order (for example, 101, 102, 103, 104, 201, 202, and so on), there is a danger of the sampling interval coinciding with the number of apartments on a floor or some multiple thereof. Then the samples might include only northwest-corner apartments or only apartments near the elevator. If these types of apartments have some other particular characteristic in common (for example, higher rent), the sample will be biased. The same danger would appear in a systematic sample of houses in a subdivision arranged with the same number of houses on a block.

In considering a systematic sample from a list, then, you should carefully examine the nature of that list. If the elements are arranged in any particular order, you should figure out whether that order will bias the sample to be selected, then you should take steps to counteract any possible bias (for example, take a simple random sample from cyclical portions).

systematic sampling A type of probability sampling in which every *k*th unit in a list is selected in the sample—for example, every 25th student in the college directory of students.

sampling interval The standard distance (k) between elements selected from a population for a sample.

sampling ratio The proportion of elements in the population that are selected to be in a sample.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Usually, however, systematic sampling is superior to simple random sampling, in convenience if nothing else. Problems in the ordering of elements in the sampling frame can usually be remedied quite easily.

Stratified Sampling

So far we've discussed two methods of sample selection from a list: random and systematic. **Stratification** is not an alternative to these methods; rather, it represents a possible modification of their use.

Simple random sampling and systematic sampling both ensure a degree of representativeness and permit an estimate of the error present. Stratified sampling is a method for obtaining a greater degree of representativeness by decreasing the probable sampling error. To understand this method, we must return briefly to the basic theory of sampling distribution.

Recall that sampling error is reduced by two factors in the sample design. First, a large sample produces a smaller sampling error than does a small sample. Second, a homogeneous population produces samples with smaller sampling errors than does a heterogeneous population. If 99 percent of the population agrees with a certain statement, it's extremely unlikely that any probability sample will greatly misrepresent the extent of agreement. If the population is split 50–50 on the statement, then the sampling error will be much greater.

Stratified sampling is based on this second factor in sampling theory. Rather than selecting a sample from the total population at large, the researcher ensures that appropriate numbers of elements are drawn from homogeneous subsets of that population. To get a stratified sample of university students, for example, you would first organize your population by college class and then draw appropriate numbers of freshmen, sophomores, juniors, and seniors. In a nonstratified sample, representation by class would be subject to the same sampling error as would other variables. In a sample stratified by class, the sampling error on this variable is reduced to zero.

More-complex stratification methods are also possible. In addition to stratifying by class, you might also stratify by gender, by GPA, and so forth. In this fashion, you might be able to ensure that your sample would contain the proper numbers of male sophomores with a 3.5 GPA, of female sophomores with a 4.0 GPA, and so forth.

The ultimate function of stratification, then, is to organize the population into homogeneous subsets (with heterogeneity between subsets) and to select the appropriate number of elements from each. To the extent that the subsets are homogeneous on the stratification variables, they may be homogeneous on other variables as well. Because age is related to college class, a sample stratified by class will also be more representative in terms of age than an unstratified sample. Because occupational aspirations still seem to be related to gender, a sample stratified by gender will be more representative in terms of occupational aspirations.

The choice of stratification variables typically depends on what variables are available. Gender can often be determined in a list of names. University lists are typically arranged by class. Lists of faculty members may indicate their departmental affiliation. Government agency files may be arranged by geographic region. Voter registration lists are arranged according to precinct.

In selecting stratification variables from among those available, however, you should be concerned primarily with those that are presumably related to variables you want to represent accurately. Because gender is related to many variables and is often available for stratification, it is often used. Education is related to many variables, but it is often not available for stratification. Geographic location within a city, state, or nation is related to many things. Within a city, stratification by geographic location usually increases representativeness in social class, ethnic group, and so forth. Within a nation, it increases representativeness in a broad range of attitudes as well as in social class and ethnicity.

When you're working with a simple list of all elements in the population, two methods of stratification predominate. In one method, you

stratification The grouping of the units composing a population into homogeneous groups (or strata) before sampling. This procedure, which may be used in conjunction with simple random, systematic, or cluster sampling, improves the representativeness of a sample, at least in terms of the variables used for stratification.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

sort the population elements into discrete groups based on whatever stratification variables are being used. On the basis of the relative proportion of the population represented by a given group, you select—randomly or systematically—several elements from that group constituting the same proportion of your desired sample size. For example, if sophomore men with a 4.0 GPA compose 1 percent of the student population and you desire a sample of 1,000 students, you would select 10 sophomore men with a 4.0 average.

The other method is to group students as described and then put those groups together in a continuous list, beginning with all male freshmen with a 4.0 average and ending with all female seniors with a 1.0 or below. You would then select a systematic sample, with a random start, from the entire list. Given the arrangement of the list, a systematic sample would select proper numbers (within an error range of 1 or 2) from each subgroup. (*Note:* A simple random sample drawn from such a composite list would cancel out the stratification.)

Figure 7-12 offers an illustration of stratified, systematic sampling. As you can see, we lined up our micropopulation according to gender and race. Then, beginning with a random start of 3, we've taken every tenth person thereafter, resulting in a list of 3, 13, 23, . . . 93.

Stratified sampling ensures the proper representation of the stratification variables; this, in turn, enhances the representation of other variables related to them. Taken as a whole, then, a stratified sample is more likely than a simple random sample to be more representative on several variables. Although the simple random sample is still regarded as somewhat sacred, it should now be clear that you can often do better.

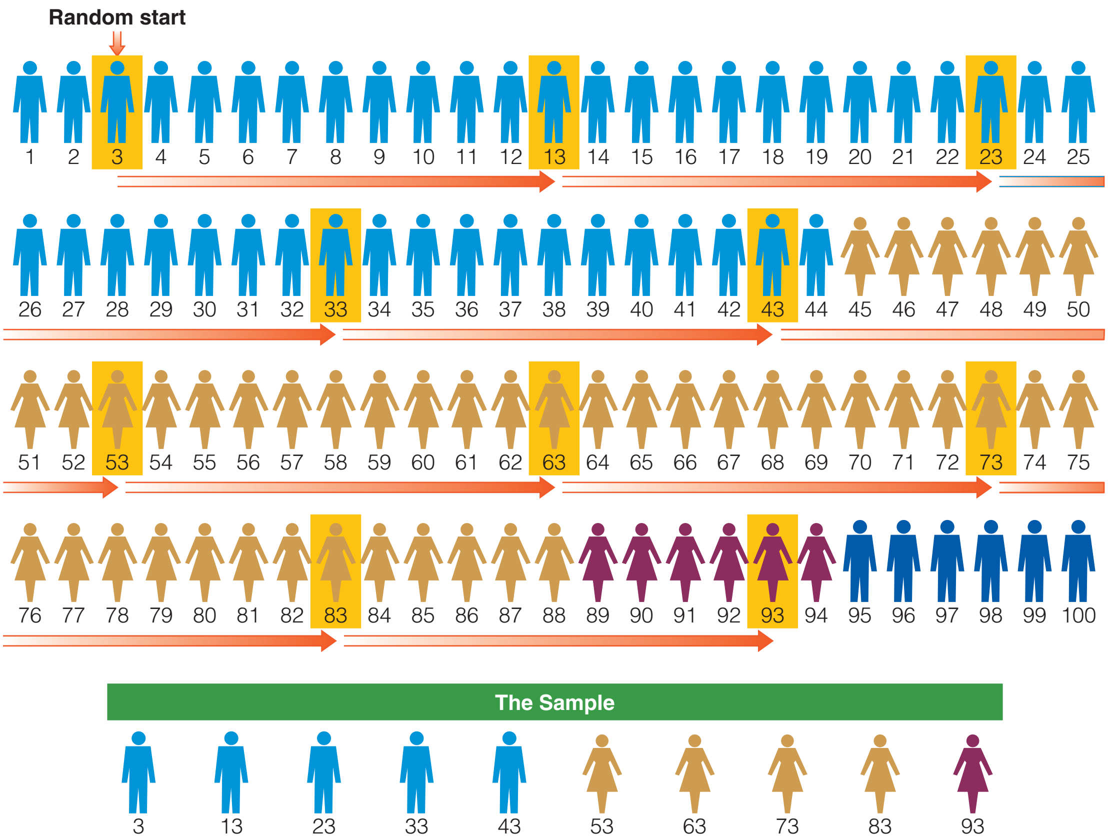


FIGURE 7-12

A Stratified, Systematic Sample with a Random Start. A stratified, systematic sample involves two stages. First the members of the population are gathered into homogeneous strata; this simple example merely uses gender and race as stratification variables, but more could be used. Then every *k*th (in this case, every tenth) person in the stratified arrangement is selected for the sample.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Implicit Stratification in Systematic Sampling

I mentioned that systematic sampling can, under certain conditions, be more accurate than simple random sampling. This is the case whenever the arrangement of the list creates an implicit stratification. As already noted, if a list of university students is arranged by class, then a systematic sample provides a stratification by class, whereas a simple random sample would not.

In a study of students at the University of Hawaii, after stratification by school class, the students were arranged by their student identification numbers. At that time, student social security numbers served as their student IDs. The first three digits of the social security number indicate the state in which the number was issued. As a result, within a class, students were arranged by the state in which they were issued a social security number, providing a rough stratification by geographic origin.

An ordered list of elements, therefore, may be more useful to you than an unordered, randomized list. I've stressed this point because of the unfortunate belief that lists should be randomized before systematic sampling. Only if the arrangement presents the problems discussed earlier should the list be rearranged.

Illustration: Sampling University Students

Let's put these principles into practice by looking at an actual sampling design used to select a sample of university students. The purpose of the study was to survey, with a mailed questionnaire, a representative cross section of students attending the main campus of the University of Hawaii. The following sections describe the steps and decisions involved in selecting that sample.

Study Population and Sampling Frame

The obvious sampling frame available for use in this sample selection was the computerized file maintained by the university administration. The file contained students' names, local and permanent addresses, social security numbers, and a variety of other information such as field of study, class, age, and gender.

The computer database, however, contained information on all people who could, by any conceivable definition, be called students, many of whom seemed inappropriate to the purposes of the study. As a result, researchers needed to define the study population in a somewhat more restricted fashion. The final definition included those 15,225 day-program degree candidates registered for the fall semester on the Manoa campus of the university, including all colleges and departments, both undergraduate and graduate students, and both U.S. and foreign students. The computer program used for sampling, therefore, limited consideration to students fitting this definition.

Stratification

The sampling program also permitted stratification of students before sample selection. The researchers decided that stratification by college class would be sufficient, although the students might have been further stratified within class, if desired, by gender, college, major, and so forth.

Sample Selection

Once the students had been arranged by class, a systematic sample was selected across the entire rearranged list. The sample size for the study was initially set at 1,100. To achieve this sample, the sampling program was set for a 1/14 sampling ratio. The program generated a random number between 1 and 14; the student having that number and every 14th student thereafter was selected in the sample.

Once the sample had been selected, the computer was instructed to print students' names and mailing addresses on self-adhesive mailing labels. These labels were then simply transferred to envelopes for mailing the questionnaires.

Sample Modification

This initial design of the sample had to be modified. Before the mailing of questionnaires, the researchers discovered that unexpected expenses in the production of the questionnaires made it impossible to cover the costs of mailing to all 1,100 students. As a result, one-third of the

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

mailing labels were systematically selected (with a random start) for exclusion from the sample. The final sample for the study was thereby reduced to 733 students.

I mention this modification in order to illustrate the frequent need to alter a study plan in midstream. Because the excluded students were systematically omitted from the initial systematic sample, the remaining 733 students could still be taken to reasonably represent the study population. The reduction in sample size did, of course, increase the range of sampling error.

Multistage Cluster Sampling

The preceding sections have dealt with reasonably simple procedures for sampling from lists of elements. Such a situation is ideal. Unfortunately, however, much interesting social research requires the selection of samples from populations that cannot easily be listed for sampling purposes: the population of a city, state, or nation; all university students in the United States; and so forth. In such cases, the sample design must be much more complex. Such a design typically involves the initial sampling of *clusters* (groups of elements), followed by the selection of elements within each of the selected clusters.

Cluster sampling may be used when it's either impossible or impractical to compile an exhaustive list of the elements composing the target population, such as all church members in the United States. Often, however, the population elements are already grouped into subpopulations, and a list of those subpopulations either exists or can be created practically. For example, church members in the United States belong to discrete churches, which are either listed or could be. Following a cluster-sample format, then, researchers would sample the list of churches in some manner (for example, a stratified, systematic sample). Next, they would obtain lists of members from each of the selected churches. Each of the lists would then be sampled, to provide samples of church members for study.

Another typical situation concerns sampling among population areas such as a city. Although there is no single list of a city's population, citizens reside on discrete city blocks or census blocks. Researchers can therefore select a sample of blocks initially, create a list of people living on each of the selected blocks, and take a subsample of the people on each block.

In a more complex design, researchers might (1) sample blocks, (2) list the households on each selected block, (3) sample the households, (4) list the people residing in each household, and (5) sample the people within each selected household. This multistage sample design leads ultimately to a selection of a sample of individuals but does not require the initial listing of all individuals in the city's population.

Multistage cluster sampling, then, involves the repetition of two basic steps: listing and sampling. The list of primary sampling units (churches, blocks) is compiled and, perhaps, stratified for sampling. Then a sample of those units is selected. The selected primary sampling units are then listed and perhaps stratified. The list of secondary sampling units is then sampled, and so forth.

The listing of households on even the selected blocks is, of course, a labor-intensive and costly activity—one of the elements making face-to-face household surveys quite expensive. Vincent Iannacchione, Jennifer Staab, and David Redden (2003) reported some initial success using postal mailing lists for this purpose, and more researchers are considering them, as we've seen. Although the lists are not perfect, they may be close enough to warrant the significant savings in cost.

Multistage cluster sampling makes possible those studies that would otherwise be impossible. Specific research circumstances often call for special designs, as demonstrated in "Issues and Insights: Sampling Iran."

cluster sampling A multistage sampling in which natural groups (clusters) are sampled initially, with the members of each selected group being subsampled afterward. For example, you might select a sample of U.S. colleges and universities from a directory, get lists of the students at all the selected schools, then draw samples of students from each.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Issues and Insights

Sampling Iran

Whereas most of the examples given in this textbook are taken from their country of origin (the United States), the basic methods of sampling would apply in other national settings as well. At the same time, researchers may need to make modifications appropriate to local conditions. In selecting a national sample of Iran, for example, Hamid Abdollahyan and Taghi Azadarmaki (2000: 21) from the University of Tehran began by stratifying the nation on the basis of cultural differences, dividing the country into nine cultural zones as follows:

1. Tehran
2. Central region including Isfahan, Arak, Qum, Yazd, and Kerman
3. The southern provinces including Hormozgan, Khuzistan, Bushehr, and Fars
4. The marginal western region including Lorestan, Charmahal and Bakhtiari, and Kogiluyeh and Eelam
5. The western provinces including western and eastern Azarbaijan, Zanjan, Ghazvin, and Ardebil
6. The eastern provinces including Khorasan and Semnan
7. The northern provinces including Gilan, Mazandran, and Golestan
8. Systan
9. Kurdistan

Within each of these cultural areas, the researchers selected samples of census blocks and, on each selected block, a sample of households. Their sample design made provisions for getting the proper numbers of men and women as respondents within households and provisions for replacing the households where no one was at home.

Though the United States and Iran are politically and culturally quite different, the sampling methods appropriate for selecting a representative sample of populations are the same. Later in this chapter, when you review a detailed description of sampling the household population of an American city, you will find it strikingly similar to the methods used in Iran by Abdollahyan and Azadarmaki.

Source: Hamid Abdollahyan and Taghi Azadarmaki. 2000. "Sampling Design in a Survey Research: The Sampling Practice in Iran." Paper presented at the meeting of the American Sociological Association, Washington, DC, August 12–16.

Multistage Designs and Sampling Error

Although cluster sampling is highly efficient, the price of that efficiency is a less-accurate sample. A simple random sample drawn from a population list is subject to a single sampling error, but a two-stage cluster sample is subject to two sampling errors. First, the initial sample of clusters will represent the population of clusters only within a range of sampling error. Second, the sample of elements selected within a given cluster will represent all the elements in that cluster only within a range of sampling error. Thus, for example, a researcher runs a certain risk of selecting a sample of disproportionately wealthy city blocks, plus a sample of disproportionately wealthy households within those blocks. The best solution to this problem lies in the number of clusters selected initially and the number of elements within each cluster.

Typically, researchers are restricted to a total sample size; for example, you may be limited to conducting 2,000 interviews in a city. Given this broad limitation, however, you have several options in designing your cluster sample. At the extremes, you could choose one cluster and select 2,000 elements within that cluster, or you could select 2,000 clusters with one element selected within each. Of course, neither approach is advisable, but a broad range of choices lies between them. Fortunately, the logic of sampling distributions provides a general guideline for this task.

Recall that sampling error is reduced by two factors: an increase in the sample size and increased homogeneity of the elements being sampled. These factors operate at each level of a multistage sample design. A sample of clusters will best represent all clusters if a large number are selected and if all clusters are very much alike. A sample of elements will best represent all elements in a given cluster if a large number are selected from the cluster and if all the elements in the cluster are very much alike.

With a given total sample size, however, if the number of clusters is increased, the number of elements within a cluster must be decreased, and vice versa. In the first case, the representativeness of the clusters is increased at the expense of more poorly representing the elements composing each cluster. Fortunately, homogeneity can be used to ease this dilemma.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Typically, the elements composing a given natural cluster within a population are more homogeneous than are all elements composing the total population. The members of a given church are more alike than are members of all churches; the residents of a given city block are more alike than are the residents of a whole city. As a result, relatively few elements may be needed to represent a given natural cluster adequately, although a larger number of clusters may be needed to adequately represent the diversity found among the clusters. This fact is most clearly seen in the extreme case of very different clusters composed of identical elements within each. In such a situation, a large number of clusters would adequately represent all its members. Although this extreme situation never exists in reality, it's closer to the truth in most cases than its opposite: identical clusters composed of grossly divergent elements.

The general guideline for cluster design, then, is to maximize the number of clusters selected while decreasing the number of elements within each cluster. However, this scientific guideline must be balanced against an administrative constraint. The efficiency of cluster sampling is based on the ability to minimize the listing of population elements. By initially selecting clusters, you need only list the elements composing the selected clusters, not all elements in the entire population. Increasing the number of clusters, however, goes directly against this efficiency factor. A small number of clusters may be listed more quickly and more cheaply than a large number. (Remember that all the elements in a selected cluster must be listed even if only a few are to be chosen in the sample.)

The final sample design will reflect these two constraints. In effect, you'll probably select as many clusters as you can afford. Lest this issue be left too open-ended at this point, here is one general guideline. Population researchers conventionally aim at the selection of 5 households per census block. If a total of 2,000 households are to be interviewed, you would aim at 400 blocks with 5 household interviews in each. Figure 7-13 presents an overview of this process.

Before turning to other, more-detailed procedures available to cluster sampling, let me reiterate that this method almost inevitably involves a loss of accuracy. The manner in which this appears, however, is somewhat complex. First, as noted earlier, a multistage sample design is subject to a sampling error at each stage. Because the sample size is necessarily smaller at each stage than the total sample size, the sampling error at each stage will be greater than would be the case for a single-stage random sample of elements. Second, sampling error is estimated on the basis of observed variance among the sample elements. When those elements are drawn from among relatively homogeneous clusters, the estimated sampling error will be too optimistic and must be corrected in light of the cluster sample design.

Stratification in Multistage Cluster Sampling

Thus far, we've looked at cluster sampling as though a simple random sample were selected at each stage of the design. In fact, stratification techniques can be used to refine and improve the sample being selected.

The basic options here are essentially the same as those in single-stage sampling from a list. In selecting a national sample of churches, for example, you might initially stratify your list of churches by denomination, geographic region, size, rural or urban location, and perhaps some measure of social class.

Once the primary sampling units (churches, blocks) have been grouped according to the relevant, available stratification variables, either simple random or systematic sampling can be used to select the sample. You might select a specified number of units from each group, or stratum, or you might arrange the stratified clusters in a continuous list and systematically sample that list.

To the extent that clusters are combined into homogeneous strata, the sampling error at this stage will be reduced. The primary goal of stratification, as before, is homogeneity.

Stratification could, of course, take place at each level of sampling. The elements listed within a selected cluster might be stratified before the next stage of sampling. Typically, however, this is not done. (Recall the assumption of relative homogeneity within clusters.)

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

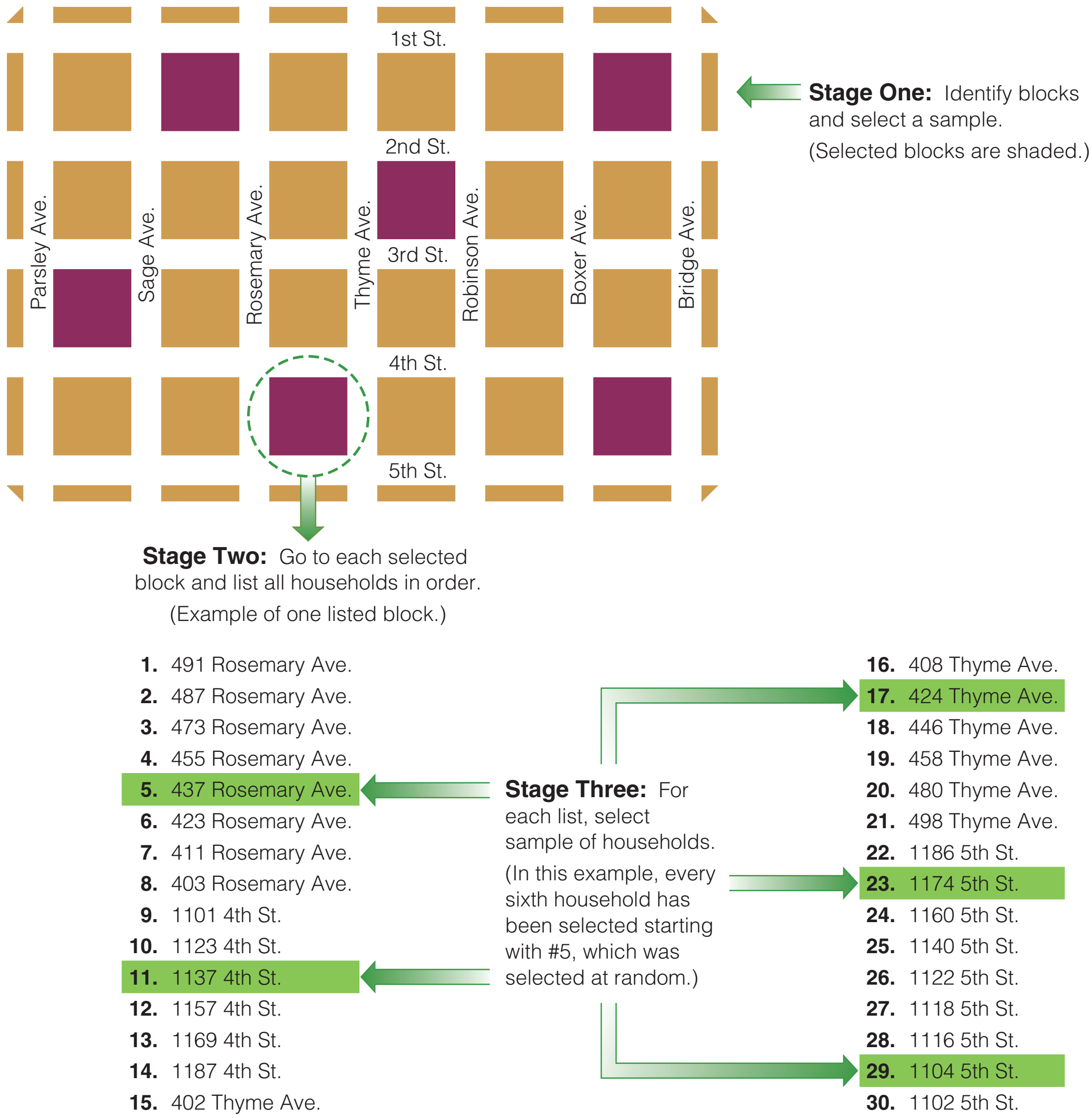


FIGURE 7-13

Multistage Cluster Sampling. In multistage cluster sampling, we begin by selecting a sample of the clusters (in this case, city blocks). Then, we make a list of the elements (households, in this case) and select a sample of elements from each of the selected clusters.

Probability Proportionate to Size (PPS) Sampling

This section introduces you to a more sophisticated form of cluster sampling, one that is used in many large-scale survey-sampling projects. In the preceding discussion, I talked about selecting a random or systematic sample of clusters and then a random or systematic sample of elements within each cluster selected. Notice that this produces an overall sampling scheme in which every element in the whole population has the same probability of selection.

Let's say we're selecting households within a city. If there are 1,000 city blocks and we initially select a sample of 100, that means that each block has a 100/1,000 or 0.1 chance of being selected. If we next select 1 household in 10 from those residing on the selected blocks, each household has a 0.1 chance of selection within

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

its block. To calculate the overall probability of a household being selected, we simply multiply the probabilities at the individual steps in sampling. That is, each household has a 1/10 chance of its block being selected and a 1/10 chance of that specific household being selected if the block is one of those chosen. Each household, in this case, has a $1/10 \times 1/10 = 1/100$ chance of selection overall. Because each household would have the same chance of selection, the sample so selected should be representative of all households in the city.

There are dangers in this procedure, however. In particular, the variation in the size of blocks (measured in numbers of households) presents a problem. Let's suppose that half the city's population resides in 10 densely packed blocks filled with high-rise apartment buildings, and suppose that the rest of the population lives in single-family dwellings spread out over the remaining 900 blocks. When we first select our sample of 1/10 of the blocks, it's quite possible that we'll miss all of the 10 densely packed high-rise blocks. No matter what happens in the second stage of sampling, our final sample of households will be grossly unrepresentative of the city, comprising only single-family dwellings.

Whenever the clusters sampled are of greatly differing sizes, it's appropriate to use a modified sampling design called **PPS (probability proportionate to size)**. This design guards against the problem I've just described and still produces a final sample in which each element has the same chance of selection.

As the name suggests, each cluster is given a chance of selection proportionate to its size. Thus, a city block with 200 households has twice the chance of selection as one with only 100 households. Within each cluster, however, a fixed number of elements is selected, say, 5 households per block. Notice how this procedure results in each household having the same probability of selection overall.

Let's look at households of two different city blocks. Block A has 100 households, whereas Block B has only 10. In PPS sampling, we would give Block A 10 times as good a chance of being selected as Block B. So if, in the overall sample design, Block A has a 1/20 chance of being selected, that means Block B would only have a 1/200 chance. Notice that this means that all the households on Block A would have a 1/20 chance of having their block selected, whereas Block B households would have only a 1/200 chance.

If Block A is selected and we're taking 5 households from each selected block, then the households on Block A have a 5/100 chance of being selected for the block's sample. Because we can multiply probabilities in a case like this, we see that every household on Block A had an overall chance of selection equal to $1/20 \times 5/100 = 5/2{,}000 = 1/400$.

If Block B happens to be selected, on the other hand, its households stand a much better chance of being among the 5 chosen there: 5/10. When this is combined with their relatively poorer chance of having their block selected in the first place, however, they end up with the same chance of selection as those on Block A: $1/200 \times 5/10 = 5/2{,}000 = 1/400$.

Further refinements to this design make it a very efficient and effective method for selecting large cluster samples. For now, however, it's enough to understand the basic logic involved.

Disproportionate Sampling and Weighting

Ultimately, a probability sample is representative of a population if all elements in the population have an equal chance of selection for that sample. Thus, in each of the preceding discussions, we've noted that the various sampling procedures result in an equal chance of selection—even though the ultimate selection probability is the product of several partial probabilities.

More generally, however, a probability sample is one in which each population element has a known nonzero probability of selection—even though different elements may have different probabilities. If controlled probability-sampling procedures have been used, any such sample may be representative of the population from which it is drawn if each sample element is

PPS (probability proportionate to size) This refers to a type of multistage cluster sample in which clusters are selected, not with equal probabilities (see *EPSEM*) but with probabilities proportionate to their sizes—as measured by the number of units to be subsampled.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

assigned a weight equal to the inverse of its probability of selection. Thus, where all sample elements have had the same chance of selection, each is given the same weight: 1. This is called a *self-weighting sample*.

Sometimes it's appropriate to give some cases more weight than others, a process called **weighting**. Disproportionate sampling and weighting come into play in two basic ways. First, you may sample subpopulations disproportionately to ensure sufficient numbers of cases from each for analysis. For example, a given city may have a suburban area containing one-fourth of its total population. Yet you might be especially interested in a detailed analysis of households in that area and may feel that one-fourth of this total sample size would be too few. As a result, you might decide to select the same number of households from the suburban area as from the remainder of the city. Households in the suburban area, then, are given a disproportionately better chance of selection than those located elsewhere in the city.

As long as you analyze the two area samples separately or comparatively, you need not worry about the differential sampling. If you want to combine the two samples to create a composite picture of the entire city, however, you must take the disproportionate sampling into account. If n is the number of households selected from each area, then the households in the suburban area had a chance of selection equal to n divided by one-fourth of the total city population. Because the total city population and the sample size are the same for both areas, the suburban-area households should be given a weight of $1/4n$, and the remaining households should be given a weight of $3/4n$. This weighting procedure could be simplified by merely giving a weight of 3 to each of the households selected outside the suburban area. (This procedure gives a proportionate representation to each sample element. The population figure would have to be included in the weighting if population estimates were desired.)

weighting Assigning different weights to cases that were selected into a sample with different probabilities of selection. In the simplest scenario, each case is given a weight equal to the inverse of its probability of selection. When all cases have the same chance of selection, no weighting is necessary.

Here's an example of the problems that can be created when disproportionate sampling is not accompanied by a weighting scheme. When the *Harvard Business Review* decided to survey its subscribers on the issue of sexual harassment at work, it seemed appropriate to oversample women because female subscribers were vastly outnumbered by male subscribers. Here's how G. C. Collins and Timothy Blodgett explained the matter:

> *We also skewed the sample another way: to ensure a representative response from women, we mailed a questionnaire to virtually every female subscriber, for a male/female ratio of 68% to 32%. This bias resulted in a response of 52% male and 44% female (and 4% who gave no indication of gender)—compared to HBR's U.S. subscriber proportion of 93% male and 7% female.*
>
> *(1981: 78)*

Notice a couple of things in this quotation. First, it would be nice to know a little more about what "virtually every female" means. Evidently, the authors of the study didn't send questionnaires to all female subscribers, but there's no indication of who was omitted and why. Second, they didn't use the term *representative* in its normal social science usage. What they mean, of course, is that they wanted to get a substantial or "large enough" response from women, and oversampling is a perfectly acceptable way of accomplishing that.

By sampling more women than a straightforward probability sample would have produced, the authors were able to "select" enough women (812) to compare with the men (960). Thus, when they report, for example, that 32 percent of the women and 66 percent of the men agree that "the amount of sexual harassment at work is greatly exaggerated," we know that the female response is based on a substantial number of cases. That's good. There are problems, however.

To begin with, subscriber surveys are always problematic. In this case, the best the researchers can hope to talk about is "what subscribers to *Harvard Business Review* think." In a loose way, it might make sense to think of that population as representing the more sophisticated portion of corporate management. Unfortunately, the overall response rate was 25 percent. Although that's quite good for subscriber surveys, it's a low response rate in terms of generalizing from probability samples.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Beyond that, however, the disproportionate sample design creates a further problem. When the authors state that 73 percent of respondents favor company policies against harassment (Collins and Blodgett 1981: 78), that figure is undoubtedly too high, because the sample contains a disproportionately high percentage of women, who are more likely to favor such policies. Further, when the researchers report that top managers are more likely to feel that claims of sexual harassment are exaggerated than are middle- and lower-level managers (1981: 81), that finding is also suspect. As the researchers report, women are disproportionately represented in lower management. That alone might account for the apparent differences among levels of management regarding harassment. In short, the failure to take account of the oversampling of women confounds all survey results that do not separate the findings by gender. The solution to this problem would have been to weight the responses by gender, as described earlier in this section.

In recent election-campaign polls, survey weighting has become a central topic, as some polling agencies weight their results on the basis of party affiliation and other variables, whereas others do not. Weighting in this instance involves assumptions regarding the differential participation of Republicans and Democrats in opinion polls and on Election Day—plus a determination of how many Republicans and Democrats there are. This will likely remain a topic of debate among pollsters and politicians in the years to come.

Probability Sampling in Review

Much of this chapter has been devoted to the key sampling method used in controlled survey research: probability sampling. In each of the variations examined, we've seen that elements are chosen for study from a population on a basis of random selection with known nonzero probabilities.

Depending on the field situation, probability sampling can be either very simple or extremely difficult, time consuming, and expensive. Whatever the situation, however, it remains the most effective method for the selection of study elements. There are two reasons for that.

First, probability sampling avoids researchers' conscious or subconscious biases in element selection. If all elements in the population have an equal (or unequal and subsequently weighted) chance of selection, there is an excellent chance that the sample so selected will closely represent the population of all elements.

Second, probability sampling permits estimates of sampling error. Although no probability sample will be perfectly representative in all respects, controlled selection methods permit the researcher to estimate the degree of expected error.

In this lengthy chapter, we've taken on a basic issue in much social research: selecting observations that will tell us something more general than the specifics we've actually observed. This issue confronts field researchers, who face more action and more actors than they can observe and record fully, as well as political pollsters who want to predict an election but can't interview all voters. As we proceed through the book, we'll see in greater detail how social researchers have found ways to deal with this issue.

The Ethics of Sampling

The key purpose of the sampling techniques discussed in this chapter is to allow researchers to make relatively few observations but gain an accurate picture of a large population. Quantitative studies using probability sampling should result in a statistical profile, based on the sample, that closely mirrors the profile that would have been gained from observing the whole population. In addition to using legitimate sampling techniques, researchers should be careful to point out the possibility of errors: sampling error, flaws in the sampling frame, nonresponse error, or anything else that might make the results misleading.

Sometimes, more typically in qualitative studies, the purpose of sampling may be to capture the breadth of variation within a population rather than to focus on the "average" or "typical" member of that population. Although this is a legitimate and valuable approach, readers may mistake the display of differences to reflect the distribution of characteristics in the population. As such, the researcher should ensure that the reader is not misled.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

What do you think?...Revisited

Contrary to common sense, we've seen that the number of people selected in a sample, while important, is less important than how people are selected. The *Literary Digest* mailed ballots to 10 million people and received 2 million back from voters around the country. However, the people they selected for their enormous sample—auto owners and telephone subscribers—were not representative of the population in 1936, during the Great Depression. Overall, the sample population was wealthier than was the voting population at large. Because rich people are more likely than the general public to vote Republican, the *Literary Digest* tallied the voting intentions of a disproportionate number of Republicans.

The probability-sampling techniques used today allow researchers to select smaller, more representative samples. Even a couple of thousand respondents, properly selected, can accurately predict the behavior of 100 million voters.

MAIN POINTS

Introduction

- Social researchers must select observations that will allow them to generalize to people and events not observed. Often this involves sampling—selecting people to observe.
- Understanding the logic of sampling is essential to doing social research.

A Brief History of Sampling

- Sometimes you can and should select probability samples using precise statistical techniques, but at other times nonprobability techniques are more appropriate.

Nonprobability Sampling

- Nonprobability-sampling techniques include reliance on available subjects, purposive (judgmental) sampling, snowball sampling, and quota sampling. In addition, researchers studying a social group may make use of informants. Each of these techniques has its uses, but none of them ensures that the resulting sample will be representative of the population being sampled.

The Logic and Techniques of Probability Sampling

- Probability-sampling methods provide an excellent way of selecting representative samples from large, known populations. These methods counter the problems of conscious and subconscious sampling bias by giving each element in the population a known (nonzero) probability of selection.
- Random selection is often a key element in probability sampling.
- The most carefully selected sample will never provide a perfect representation of the population from which it was selected. There will always be some degree of sampling error.
- By predicting the distribution of samples with respect to the target parameter, probability-sampling methods make it possible to estimate the amount of sampling error expected in a given sample.
- The expected error in a sample is expressed in terms of confidence levels and confidence intervals.

Populations and Sampling Frames

- A sampling frame is a list or quasi-list of the members of a population. It is the resource used in the selection of a sample. A sample's representativeness depends directly on the extent to which a sampling frame contains all the members of the total population that the sample is intended to represent.
- We've seen that the proliferation of cell phones is complicating the task of listing the desired population and selecting a sample.

Types of Sampling Designs

- Several sampling designs are available to researchers.
- Simple random sampling is logically the most fundamental technique in probability sampling, but it is seldom used in practice.
- Systematic sampling involves the selection of every *k*th member from a sampling frame. This method is more practical than simple random sampling and, with a few exceptions, is functionally equivalent.
- Stratification, the process of grouping the members of a population into relatively homogeneous strata before sampling, improves the representativeness of a sample by reducing the degree of sampling error.

Multistage Cluster Sampling

- Multistage cluster sampling is a relatively complex sampling technique that is frequently used when a list of all the members of a population does not exist. Typically, researchers must balance the number of clusters and the size of each cluster to achieve a given sample size.

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s). Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.

Stratification can be used to reduce the sampling error involved in multistage cluster sampling.

- Probability proportionate to size (PPS) is a special, efficient method for multistage cluster sampling.
- If the members of a population have unequal probabilities of selection for the sample, researchers must assign weights to the different observations made in order to provide a representative picture of the total population. Basically, the weight assigned to a particular sample member should be the inverse of its probability of selection.

Probability Sampling in Review

- Probability sampling remains the most effective method for the selection of study elements because (1) it allows researchers to avoid biases in element selection and (2) it permits estimates of error.

The Ethics of Sampling

- Probability sampling always carries a risk of error; researchers must inform readers of any errors that might make results misleading.
- When nonprobability-sampling methods are used to obtain the breadth of variations in a population, researchers must take care not to mislead readers into confusing variations with what's typical in the population.

KEY TERMS

cluster sampling
confidence interval
confidence level
element
EPSEM (equal probability of selection method)
informant
nonprobability sampling
parameter
population
PPS (probability proportionate to size)
probability sampling
purposive sampling
quota sampling
random selection
representativeness
sampling error
sampling frame
sampling interval
sampling ratio
sampling unit
simple random sampling
snowball sampling
statistic
stratification
study population
systematic sampling
weighting

PROPOSING SOCIAL RESEARCH: SAMPLING

In this portion of the proposal, you'll describe how you'll select from among all the possible observations you might make. Depending on the data-collection method you plan to employ, either probability or nonprobability sampling may be appropriate for your study. Similarly, this aspect of your proposal may involve the sampling of subjects or informants, or it could involve the sampling of corporations, cities, books, and so forth.

Your proposal, then, must specify what units you'll be sampling among, the data you'll use for purposes of your sample selection (your sampling frame, for instance), and the actual sampling methods you plan to use.

REVIEW QUESTIONS

1. Review the discussion of the 1948 Gallup poll that predicted that Thomas Dewey would defeat Harry Truman for president. What are some ways Gallup could have modified his quota-sampling design to avoid the error?
2. Using Appendix B of this book, select a simple random sample of 10 numbers in the range from 1 to 9,876. What is each step in the process?
3. What are the steps involved in selecting a multistage cluster sample of students taking first-year English in U.S. colleges and universities?
4. In Chapter 9, we'll discuss surveys conducted on the Internet. Can you anticipate possible problems concerning sampling frames, representativeness, and the like? Do you see any solutions?

Copyright 2021 Cengage Learning. All Rights Reserved. May not be copied, scanned, or duplicated, in whole or in part. Due to electronic rights, some third party content may be suppressed from the eBook and/or eChapter(s).
Editorial review has deemed that any suppressed content does not materially affect the overall learning experience. Cengage Learning reserves the right to remove additional content at any time if subsequent rights restrictions require it.