CHAPTER 5

Sampling

Sample Planning

Define Sample Components and the Population
Evaluate Generalizability
Assess the Diversity of the Population
Consider a Census

Sampling Methods

Probability Sampling Methods
Simple Random Sampling
Systematic Random Sampling
Stratified Random Sampling
Cluster Sampling
Probability Sampling Methods Compared

Nonprobability Sampling Methods
Availability Sampling
Quota Sampling
Purposive Sampling
Snowball Sampling
Lessons About Sample Quality
Generalizability in Qualitative Research

Sampling Distributions

Estimating Sampling Error
Determining Sample Size

Conclusions

A common technique in journalism is to put a "human face" on a story. For instance, a *Boston Globe* reporter (Abel 2008) interviewed a participant for a story about a housing program for chronically homeless people. "Burt" had worked as a welder, but alcoholism and both physical and mental health problems interfered with his plans. By the time he was 60, Burt had spent many years on the streets. Fortunately, he obtained an independent apartment through a new Massachusetts program, but even then "the lure of booze and friends from the street was strong" (Abel 2008:A14).

It is a sad story with an all-too-uncommon happy—although uncertain—ending. Together with one other such story and comments by several service staff, the article provides a persuasive rationale for the new housing program. However, we don't know whether the two participants interviewed for the story are like most program participants, most homeless persons in Boston, or most homeless persons throughout the United States—or whether they

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:25 PM Page 149

are just two people who caught the eye of this one reporter. In other words, we don't know how generalizable their stories are, and if we don't have confidence in generalizability, then the validity of this account of how the program participants became homeless is suspect. Because we don't know whether their situation is widely shared or unique, we cannot really judge what the account tells us about the social world.

In this chapter, you will learn about sampling methods, the procedures that primarily determine the generalizability of research findings. I first review the rationale for using sampling in social research and consider two circumstances when sampling is not necessary. The chapter then turns to specific sampling methods and when they are most appropriate, using examples from research on homelessness. This section is followed by a section on sampling distributions, which introduces you to the logic of statistical inference—that is, how to determine how likely it is that our sample statistics represent the population from which the sample was drawn. By the chapter's end, you should understand which questions you need to ask to evaluate the generalizability of a study as well as what choices you need to make when designing a sampling strategy. You should also realize that it is just as important to select the "right" people or objects to study as it is to ask participants the right questions.

SAMPLE PLANNING

You have encountered the problem of generalizability in each of the studies you have read about in this book. For example, Keith Hampton and Barry Wellman (1999) discussed their findings in Netville as though they could be generalized to residents of other communities; Norman Nie and Lutz Erbring (2000) generalized their Internet survey findings to the entire American adult population, and the National Geographic Society (2000) Web survey findings were generalized to the entire world. Whether we are designing a sampling strategy or evaluating someone else's findings, we have to understand how and why researchers decide to sample and what the consequences of these decisions are for the generalizability of the study's findings.

Define Sample Components and the Population

Let's say that we are designing a survey about adult homeless persons in one city. We don't have the time or resources to study the entire adult **population** of the city, even though it consists of the set of individuals or other entities to which we wish to be able to generalize our findings. Even the city of Boston, which conducts an annual census of homeless persons, does not have the resources to actually survey the homeless persons they count. So instead, we resolve to study a **sample**, a subset of this population. The individual members of this sample are called **elements**, or elementary units.

In many studies, we sample directly from the elements in the population of interest. We may survey a sample of the

Population The entire set of individuals or other entities to which study findings are to be generalized.

Sample A subset of a population that is used to study the population as a whole.

Elements The individual members of the population whose characteristics are to be measured.

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:25 PM Page 150

Sampling frame A list of all elements or other units containing the elements in a population.

Enumeration units Units that contain one or more elements and that are listed in a sampling frame.

Sampling units Units listed at each stage of a multistage sampling design.

entire population of students in a school, based on a list obtained from the registrar's office. This list, from which the elements of the population are selected, is termed the **sampling frame.** The students who are selected and interviewed from that list are the elements.

In some studies, the entities that can be reached easily are not the same as the elements from which we want information, but they include those elements. For example, we may have a list of households but not a list of the entire population of a town, even though the adults are the elements that we actually want to sample. In this situation, we could draw a sample of households so that we can identify the adult individuals in these households. The households are termed **enumeration units**, and the adults in the households are the elements (Levy & Lemeshow 1999:13–14).

Sometimes, the individuals or other entities from which we collect information are not actually the elements in our study. For example, a researcher might sample schools for a survey about educational practices and then interview a sample of teachers in each sampled school to obtain the information about educational practices. Both the schools and the teachers are termed **sampling units**, because we sample from both (Levy & Lemeshow 1999:22). The schools are selected in the first stage of the sample, so they are the *primary sampling units* (in this case, they are also the elements in the study). The teachers are *secondary sampling units* (but they are not elements, because they are used to provide information about the entire school) (see Exhibit 5.1).

It is important to know exactly what population a sample can represent when you select or evaluate sample components. In a survey of "adult Americans," the general population may reasonably be construed as all residents of the United States who are at least 21 years old. But always be alert to ways in which the population may have been narrowed by the sample selection procedures. For example, perhaps only English-speaking residents of the United States were surveyed. The population for a study is the aggregation of elements that we actually focus on and sample from, not some larger aggregation that we really wish we could have studied.

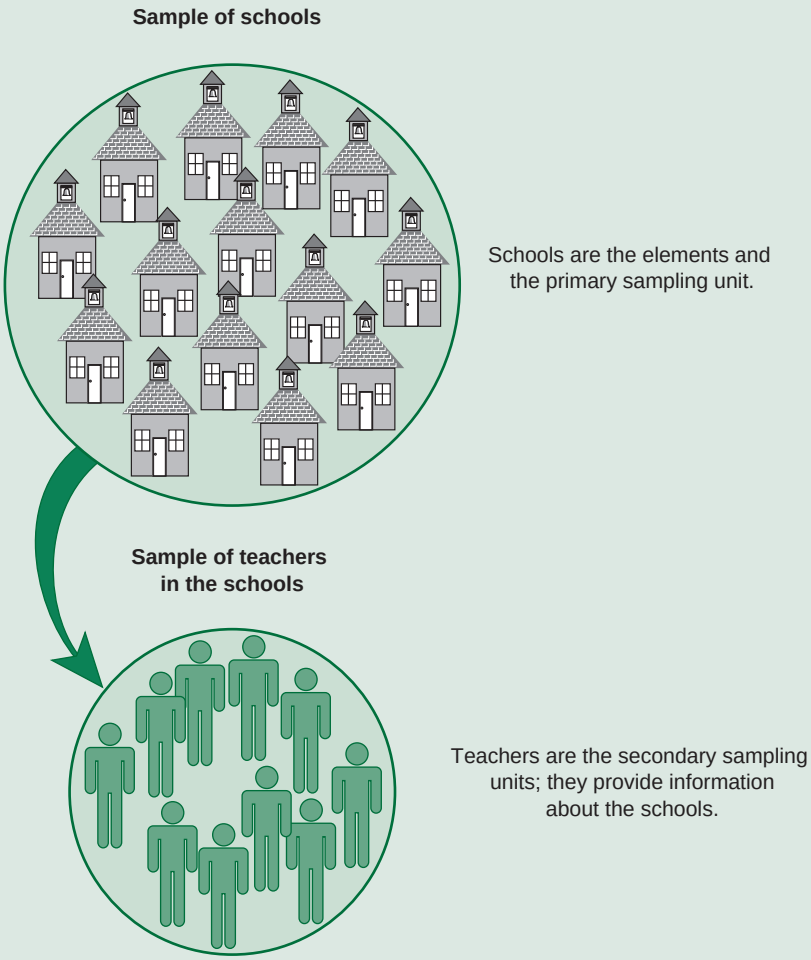
Some populations, such as the homeless, are not identified by a simple criterion such as a geographic boundary or an organizational membership. Clear definition of such a population is difficult but quite necessary. Anyone should be able to determine just what population was actually studied. However, studies of homeless persons in the early 1980s "did not propose definitions, did not use screening questions to be sure that the people they interviewed were indeed homeless, and did not make major efforts to cover the universe of homeless people" (Burt 1996:15). (Perhaps just homeless persons in one shelter were studied.) The result was a "collection of studies that could not be compared" (Burt 1996:15). Several studies of homeless persons in urban areas addressed the problem by employing a more explicit definition of the population: "people who had no home or permanent place to stay of their own (meaning they rented or owned it themselves) and no regular arrangement to stay at someone else's place" (Burt 1996:18).

Even this more explicit definition still leaves some questions unanswered: What is a "regular arrangement"? How permanent does a "permanent place" have to be? In a study of

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:26 PM Page 151

EXHIBIT 5.1 Sample Components in a Two-Stage Study



homeless persons in Chicago, Michael Sosin, Paul Colson, and Susan Grossman (1988) answered these questions in their definition of the population of interest:

> We define the homeless as: those current[ly] residing for at least one day but for less than fourteen with a friend or relative, not paying rent, and not sure that the length of stay will surpass fourteen days; those currently residing in a shelter, whether overnight or transitional; those currently without normal, acceptable shelter arrangements and thus sleeping on the street, in doorways, in abandoned buildings, in cars, in subway or bus stations, in alleys, and so forth; those residing in a treatment center for the indigent who have lived at the facility for less than 90 days and who claim that they have no place to go, when released. (p. 22)

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:26 PM Page 152

This definition reflects accurately Sosin et al.'s concept of homelessness and allows researchers in other locations or at other times to develop procedures for studying a comparable population. The more complete and explicit the definition is of the population from which a sample was selected, the more precise our generalizations can be.

Evaluate Generalizability

Once we have defined clearly the population from which we will sample, we need to determine the scope of the generalizations we will make from our sample. Do you recall from Chapter 2 the two different meanings of generalizability?

Can the findings from a sample of the population be generalized to the population from which the sample was selected? Did Nie and Erbring's (2000) findings apply to the United States, National Geographic's (2000) to the entire world, or Wechsler et al.'s (2000) study of binge drinking to all U.S. college students? This type of generalizability was defined as *sample generalizability* in Chapter 2.

Can the findings from a study of one population be generalized to another, somewhat different population? Are e-mail users in Netville similar to those in other Ontario suburbs? In other provinces? In the United States? Are students similar to full-time employees, housewives, or other groups in their drinking patterns? Do findings from a laboratory study about alcohol effects at a small northeastern U.S. college differ from those that would be obtained at a college in the Midwest? What is the generalizability of the results from a survey of homeless persons in one city? This type of generalizability question was defined as *cross-population generalizability* in Chapter 2.

This chapter focuses attention primarily on the problem of sample generalizability: Can findings from a sample be generalized to the population from which the sample was drawn? This is really the most basic question to ask about a sample, and social research methods provide many tools with which to address it.

Sample generalizability depends on sample quality, which is determined by the amount of **sampling error**—the difference between the characteristics of a sample and the characteristics of the population from which it was selected. The larger the sampling error, the less representative the sample—and thus the less generalizable the findings. To assess sample quality when you are planning or evaluating a study, ask yourself these questions:

- From what population were the cases selected?
- What method was used to select cases from this population?
- Do the cases that were studied represent, in the aggregate, the population from which they were selected?

But researchers often project their theories onto groups or populations much larger than, or simply different from, those they have actually studied. The population to which generalizations are made in this way can be termed the **target population**—a set of elements larger than or different from the population that was sampled and to which the researcher would like to generalize any study findings. When we generalize findings to target populations, we must be somewhat speculative. We must carefully consider the validity of claims that the findings can be applied to other groups, geographic areas, cultures, or times.

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:26 PM Page 153

Because the validity of cross-population generalizations cannot be tested empirically, except by conducting more research in other settings, I will not focus much attention on this problem here. But I'll return to the problem of cross-population generalizability in Chapter 7, which addresses experimental research, and in Chapter 12, which discusses methods for studying different societies.

Sampling error Any difference between the characteristics of a sample and the characteristics of a population. The larger the sampling error, the less representative the sample.

Target population A set of elements larger than or different from the population sampled and to which the researcher would like to generalize study findings.

Assess the Diversity of the Population

Sampling is unnecessary if all the units in the population are identical. Physicists don't need to select a representative sample of atomic particles to learn about basic physical processes. They can study a single atomic particle because it is identical to every other particle of its type. Similarly, biologists don't need to sample a particular type of plant to determine whether a given chemical has toxic effects on that particular type. The idea is "If you've seen one, you've seen 'em all."

What about people? Certainly, all people are not identical (nor are other animals, in many respects). Nonetheless, if we are studying physical or psychological processes that are the same among all people, sampling is not needed to achieve generalizable findings. Psychologists and social psychologists often conduct experiments on college students to learn about processes that they think are identical across individuals. They believe that most people would have the same reactions as the college students if they experienced the same experimental conditions. Field researchers who observed group processes in a small community sometimes make the same assumption.

There is a potential problem with this assumption, however: There's no way to know for sure if the processes being studied are identical across all people. In fact, experiments can give different results depending on the type of people who are studied or the conditions for the experiment. Stanley Milgram's (1965) classic experiments on obedience to authority, which you studied in Chapter 3, illustrate this point very well. You remember that the original Milgram experiments tested the willingness of male volunteers in New Haven, Connecticut, to comply with the instructions of an authority figure to give "electric shocks" to someone else, even when these shocks seemed to harm the person receiving them. In most cases, the volunteers complied. Milgram concluded that people are very obedient to authority.

Were these results generalizable to all men, to men in the United States, or to men in New Haven? The initial experiment was repeated many times to assess the generalizability of the findings. Similar results were obtained in many replications of the Milgram experiments, that is, when the experimental conditions and subjects were similar to those studied by Milgram. Other studies showed that some groups were less likely to react so obediently. Given certain conditions, such as another "subject" in the room who refused to administer the shocks, subjects were likely to resist authority.

So what do the initial experimental results tell us about how people will react to an authoritarian movement in the real world, when conditions are not so carefully controlled? In the real social world, people may be less likely to react obediently as well. Other individuals may argue against obedience to a particular leader's commands, or people may see on TV the consequences of their actions. But alternatively, people in the real world may be even more

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:26 PM Page 154

Representative sample A sample that "looks like" the population from which it was selected in all respects that are potentially relevant to the study. The distribution of characteristics among the elements of a representative sample is the same as the distribution of those characteristics among the total population. In an unrepresentative sample, some characteristics are overrepresented or underrepresented.

obedient to authority than were the experimental subjects, for example, when they get swept up in mobs or are captivated by ideological fervor. Milgram's initial research and the many replications of it give us great insight into human behavior, in part, because they help identify the types of people and conditions to which the initial findings (lack of resistance to authority) can be generalized. But generalizing the results of single experiments is always risky, because such research often studies a small number of people who are not selected to represent any particular population.

The main point is that social scientists rarely can skirt the problem of demonstrating the generalizability of their findings. If a small sample has been studied in an experiment or a field research project, the study should be replicated in different settings or, preferably, with a **representative sample** of the population to which generalizations are sought (see Exhibit 5.2). The social world and the people in it are just too diverse to be considered "identical units." Social psychological experiments and small field studies have produced good social science, but they need to be replicated in other settings, with other subjects, to claim any generalizability. Even when we believe that we have uncovered basic social processes in a laboratory experiment or field observation, we should be very concerned with seeking confirmation in other samples and in other research.

Census Research in which information is obtained through the responses that all available members of an entire population give to questions.

Consider a Census

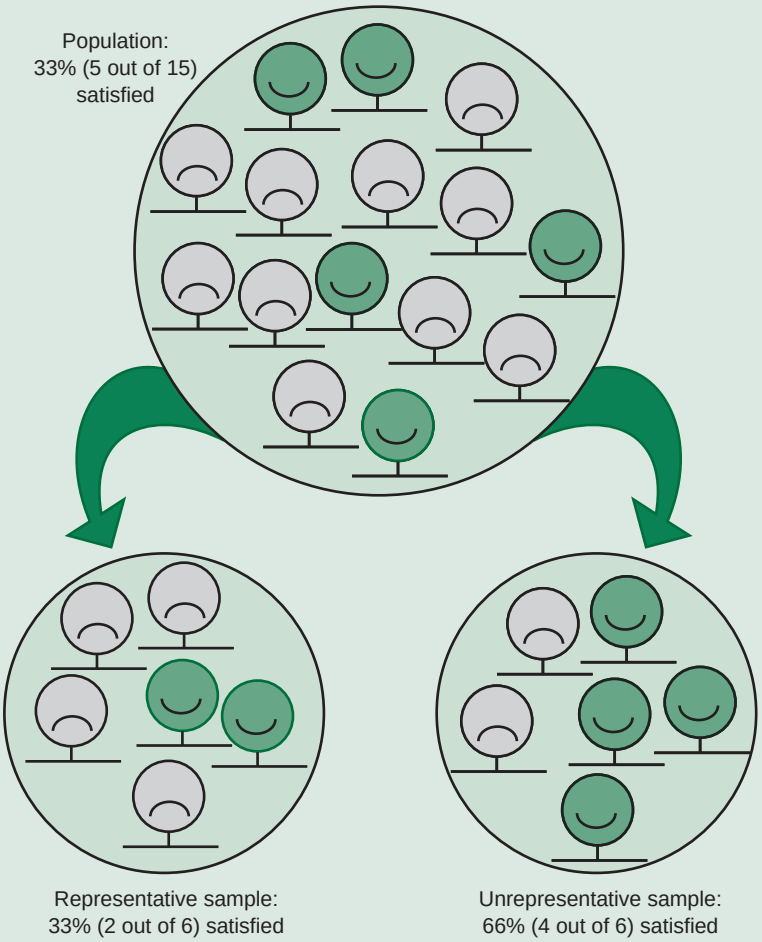
In some circumstances, it may be feasible to skirt the issue of generalizability by conducting a **census**—studying the entire population of interest—rather than drawing a sample. This is what the federal government tries to do every 10 years with the U.S. Census. Censuses also include studies of all the employees (or students) in small organizations, studies comparing all 50 states, and studies of the entire population of a particular type of organization in some area. However, in comparison with the U.S. Census and similar efforts in other countries, states, and cities, the population that is studied in these other censuses is relatively small.

The reason that social scientists don't often attempt to collect data from all the members of some large population is simply that doing so would be too expensive and time-consuming—and they can do almost as well with a sample. Some social scientists conduct research with data from the U.S. Census, but it's the government that collects the data and it's your tax dollars that pay for the effort. To conduct the 2000 Census, the Congress and the president allocated almost $4.5 billion (Prewitt 2000), and the U.S. Bureau of the Census spent 12 years doing the planning (U.S. Bureau of the Census, 2000a). For the 2010 Census, the Census Bureau is already testing new approaches, including an Internet-based response option (U.S. Bureau of the Census 2003).

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:26 PM Page 155

EXHIBIT 5.2 Representative and Unrepresentative Samples



Even if the population of interest for a survey is a small town of 20,000 or students in a university of 10,000, researchers will have to sample. The costs of surveying "just" thousands of individuals exceed by far the budgets for most research projects. In fact, not even the U.S. Bureau of the Census can afford to have everyone answer all the questions that should be covered in the census. So it draws a sample. Every household must complete a short version of the census (it had seven basic questions in 2000), and a sample consisting of one in six households must complete a long form (with 53 additional questions) (Rosenbaum 2000).

The fact is that it is hard to get people to complete a survey is another reason why survey research can be costly. Even the U.S. Bureau of the Census (1999) must make multiple efforts

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:26 PM Page 156

to increase the rate of response in spite of the federal law requiring all citizens to complete their census questionnaire. After the Census Bureau spent $167 million on publicity (Forero 2000b), two-thirds of the population returned their census questionnaire through the mail, ending a three-decade decline (U.S. Bureau of the Census 2000e). However, half a million temporary workers and up to six follow-ups were required to contact the rest of the households that did not respond by mail (U.S. Bureau of the Census 2000b, 2000c). As the U.S. 2000 Census progressed, concerns arose about underrepresentation of minority groups (Kershaw 2000), impoverished cities (Zielbauer 2000), well-to-do individuals in gated communities and luxury buildings (Langford 2000), and even college students (Abel 2000), so the Bureau conducted an even more intensive sample survey to learn about the characteristics of those who had still not responded (Anderson & Fienberg 1999; U.S. Bureau of the Census, 2000d). The number of persons missed in the census was still estimated to be between 3.2 and 6.4 million (U.S. Bureau of the Census 2001), and controversy continued over underrepresentation of some groups (Armas 2002; Holmes 2001a).

The average survey project has far less legal and financial backing, and so an adequate census is not likely to be possible. Even in Russia, which spent almost $200 million to survey its population of about 145 million, resource shortages after the collapse of the Soviet Union prevented an adequate census (Myers 2002). The census had to be postponed from 1999 to 2002 due to insufficient funds and had to rely on voluntary participation. In spite of an $8 million advertising campaign, may residents in impoverished regions refused to take part (Tavernise 2002). In Vladivostok, "many residents, angry about a recent rise in electricity prices, refused to take part. Residents on Russian Island . . . boycotted to protest dilapidated roads" (Tavernise 2002:A13).

In most survey situations, it is much better to survey only a limited number from the total population so that there are more resources for follow-up procedures that can overcome reluctance or indifference about participation. (I will give more attention to the problem of nonresponse in Chapter 8.)

SAMPLING METHODS

We can now study more systematically the features of samples that make them more or less likely to represent the population from which they are selected. The most important distinction that needs to be made about the samples is whether they are based on a probability or a nonprobability sampling method. Sampling methods that allow us to know in advance how likely it is that any element of a population will be selected for the sample are termed **probability sampling methods.** Sampling methods that do not let us know in advance the likelihood of selecting each element are termed **nonprobability sampling methods**.

Probability sampling method A sampling method that relies on a random, or chance, selection method so that the probability of selection of population elements is known.

Nonprobability sampling method Sampling method in which the probability of selection of population elements is unknown.

Probability sampling methods rely on a random, or chance, selection procedure, which is, in principle, the same as flipping a coin to decide which of two people "wins" and which one "loses." Heads and tails are equally

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:26 PM Page 157

likely to turn up in a coin toss, so both persons have an equal chance of winning. That chance, their **probability of selection,** is 1 out of 2, or .5.

Probability of selection The likelihood that an element will be selected from the population for inclusion in the sample. In a census of all elements of a population, the probability that any particular element will be selected is 1.0. If half of the elements in the population are sampled on the basis of chance (say, by tossing a coin), the probability of selection for each element is one half, or .5. As the size of the sample as a proportion of the population decreases, so does the probability of selection.

Flipping a coin is a fair way to select one of two people because the selection process harbors no systematic bias. You might win or lose the coin toss, but you know that the outcome was due simply to chance, not to bias. For the same reason, a roll of a six-sided die is a fair way to choose one of six possible outcomes (the odds of selection are 1 out of 6, or .17). Dealing out a hand after shuffling a deck of cards is a fair way to allocate sets of cards in a poker game (the odds of each person getting a particular outcome, such as a full house or a flush, are the same). Similarly, state lotteries use a random process to select winning numbers. Thus, the odds of winning a lottery, the probability of selection, are known, even though they are very much smaller (perhaps 1 out of 1 million) than the odds of winning a coin toss.

Random sampling A method of sampling that relies on a random, or chance, selection method so that every element of the sampling frame has a known probability of being selected.

There is a natural tendency to confuse the concept of **random sampling**, in which cases are selected only on the basis of chance, with a haphazard method of sampling. On first impression, "leaving things up to chance" seems to imply not exerting any control over the sampling method. But to ensure that nothing but chance influences the selection of cases, the researcher must proceed very methodically, leaving nothing to chance except the selection of the cases themselves. The researcher must follow carefully controlled procedures if a purely random process is to occur. In fact, when reading about sampling methods, do not assume that a random sample was obtained just because the researcher used a random selection method at some point in the sampling process. Look for those two particular problems; selecting elements from an incomplete list of the total population and failing to obtain an adequate response rate.

If the sampling frame is incomplete, a sample selected randomly from that list will not really be a random sample of the population. You should always consider the adequacy of the sampling frame. Even for a simple population such as a university's student body, the registrar's list is likely to be at least a bit out-of-date at any given time. For example, some students will have dropped out, but their status will not yet be officially recorded. Although you may judge the amount of error introduced in this particular situation to be negligible, the problems are greatly compounded for a larger population. The sampling frame for a city, state, or nation is always likely to be incomplete because of constant migration into and out of the area. Even unavoidable omissions from the sampling frame can bias a sample against particular groups within the population.

A very inclusive sampling frame may still yield systematic bias if many sample members cannot be contacted or refuse to participate. Nonresponse is a major hazard in survey research because **nonrespondents** are likely to differ systematically from those who take the time to participate. You should not assume that findings from a randomly selected sample will be

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:26 PM Page 158

Nonrespondents People or other entities who do not participate in a study although they are selected for the sample.

generalizable to the population from which the sample was selected if the rate of nonresponse is considerable (certainly not if it is much above 30%).

Probability Sampling Methods

Systematic bias Overrepresentation or underrepresentation of some population characteristics in a sample due to the method used to select the sample. A sample shaped by systematic sampling error is a biased sample.

Probability sampling methods are those in which the probability of selection is known and is not zero (so there is some chance of selecting each element). These methods randomly select elements and therefore have no **systematic bias**; nothing but chance determines which elements are included in the sample. This feature of probability samples makes them much more desirable than nonprobability samples, when the goal is to generalize to a larger population.

Although a random sample has no systematic bias, it will certainly have some sampling error due to chance. The probability of selecting a head is .5 in a single toss of a coin and in 20, 30, or however many tosses of a coin you like. But it is perfectly possible to toss a coin twice and get a head both times. The random "sample" of the two sides of the coin is selected in an unbiased fashion, but it still is unrepresentative. Imagine selecting randomly a sample of 10 people from a population comprising 50 men and 50 women. Just by chance, can't you imagine finding that these 10 people include 7 women and only 3 men? Fortunately, we can determine mathematically the likely degree of sampling error in an estimate based on a random sample (as we'll discuss later in this chapter)—assuming that the sample's randomness has not been destroyed by a high rate of nonresponse or by poor control over the selection process.

In general, both the size of the sample and the homogeneity (sameness) of the population affect the degree of error due to chance; the proportion of the population that the sample represents does not. To elaborate,

- *The larger the sample, the more confidence we can have in the sample's representativeness.* If we randomly pick 5 people to represent the entire population of our city, our sample is unlikely to be very representative of the entire population in terms of age, gender, race, attitudes, and so on. But if we randomly pick 100 people, the odds of having a representative sample are much better; with a random sample of 1,000, the odds become very good indeed.
- *The more homogeneous the population, the more confidence we can have in the representativeness of a sample of any particular size.* Let's say we plan to draw samples of 50 from each of two communities to estimate mean family income. One community is very diverse, with family incomes varying from $12,000 to $85,000. In the other, more homogeneous community, family incomes are concentrated in a narrow range, from $41,000 to $64,000. The estimated mean family income based on the sample from the homogeneous community is more likely to be representative than is the estimate based on the sample from the more heterogeneous community. With less variation to represent, fewer cases are needed to represent the homogeneous community.
- *The fraction of the total population that a sample contains does not affect the sample's representativeness unless that fraction is large.* We can regard any sampling fraction

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

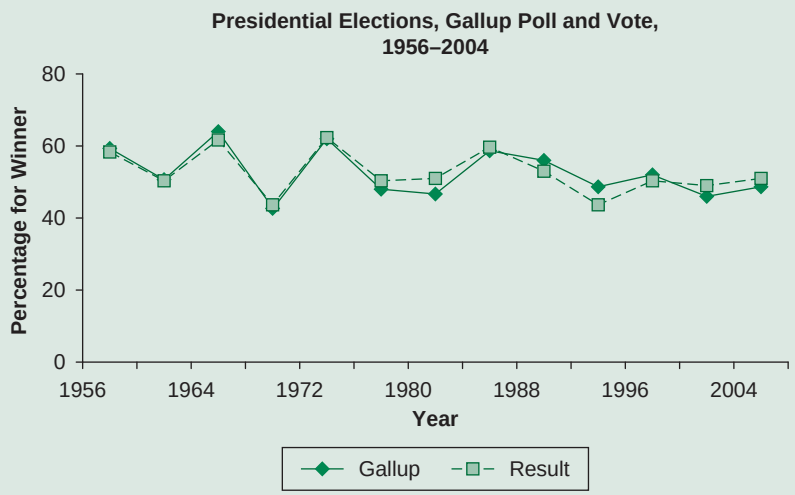
05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:27 PM Page 159

less than 2% with about the same degree of confidence (Sudman 1976:184). In fact, sample representativeness is not likely to increase much until the sampling fraction is quite a bit higher. Other things being equal, a sample of 1,000 from a population of 1 million (with a sampling fraction of 0.001, or 0.1%) is much better than a sample of 100 from a population of 10,000 (although the sampling fraction for this smaller sample is 0.01, or 1%, which is 10 times higher). The size of the samples is what makes representativeness more likely, not the proportion of the whole that the sample represents.

Polls to predict presidential election outcomes illustrate both the value of random sampling and the problems that it cannot overcome. In most presidential elections, pollsters have predicted accurately the outcomes of the actual votes by using random sampling and, these days, phone interviewing to learn for which candidate the likely voters intend to vote. Exhibit 5.3 shows how close these sample-based predictions have been in the last 13 contests. The exceptions were the 1980 and 1992 elections, when third-party candidates had an unpredicted effect. Otherwise, the small discrepancies between the votes predicted through random sampling and the actual votes can be attributed to random error.

The Gallup poll did quite well in predicting the result of the hotly contested 2000 presidential election. The final Gallup prediction was that George W. Bush would win with 48% (Al Gore was predicted to receive only 46%, while Green Party nominee Ralph Nader was predicted to secure 4%). Although the race turned out much closer, with Gore actually winning the popular vote (before losing in the electoral college), Gallup accurately noted that there appeared to have been a late-breaking trend in favor of Gore (Newport 2000). In 2004, the final Gallup prediction of 49% for Bush was within 2 percentage points of his winning total of 51% (actually, 50.77%); the "error" is partially due to the 1% of votes cast for third-party candidate Ralph Nader.

EXHIBIT 5.3 Presidential Election Outcomes: Predicted and Actual



Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:27 PM Page 160

Nevertheless, election polls have produced some major errors in prediction. The reasons for these errors illustrate some of the ways in which unintentional systematic bias can influence sample results. In 1936, a *Literary Digest* poll predicted that Alfred M. Landon would defeat President Franklin Delano Roosevelt in a landslide, but instead Roosevelt took 63% of the popular vote. The problem? The *Digest* mailed out 10 million mock ballots to people listed in telephone directories, automobile registration records, voter lists, and so on. But in 1936, during the Great Depression, only relatively wealthy people had phones and cars, and they were more likely to be Republican. Furthermore, only 2,376,523 completed ballots were returned, and a response rate of only 24% leaves much room for error. Of course, this poll was not designed as a random sample, so the appearance of systematic bias is not surprising. Gallup was able to predict the 1936 election results accurately with a randomly selected sample of just 3,000 (Bainbridge 1989:43–44).

In 1948, pollsters mistakenly predicted that Thomas E. Dewey would beat Harry S. Truman, based on the random sampling method that George Gallup had used successfully since 1934. The problem? Pollsters stopped collecting data several weeks before the election, and in those weeks, many people changed their minds (Kenney, 1987). The sample was systematically biased by underrepresenting shifts in voter sentiment just before the election.

The fast-paced 2008 presidential primary elections were also challenging for the pollsters, primarily among Democratic Party voters. In the early New Hampshire primary, polls successfully predicted Republican John McCain's winning margin of 5.5% (the polls were off by only 0.2%, on average). However, all the polls predicted that Barack Obama would win the Democratic primary by a margin of about 8 percentage points, but he lost to Hillary Clinton by 12 points (47% to 35%). In a careful review of different explanations that have been proposed for that failure, the president of the Pew Research Center, Andrew Kohut (2008:A27), concluded that the problem was that voters who are poorer, less well educated, and white and who tend to refuse to respond to surveys tend to be less favorable to blacks than other voters. These voters, who were unrepresented in the polls, were more likely to favor Clinton over Obama.

Because they do not disproportionately exclude or include particular groups within the population, random samples that are successfully implemented avoid systematic bias in the selection process. However, when some types of people are more likely to refuse to participate in surveys or are less likely to be available for interviews, systematic bias can still creep into the sampling process. In addition, random error will still influence the specific results obtained from any random sample. Different types of random samples vary in their ability to minimize random error. The four most common methods for drawing random samples are simple random sampling, systematic random sampling, stratified random sampling, and cluster sampling.

Simple random sampling A method of sampling in which every sample element is selected only on the basis of chance, through a random process.

Random number table A table containing lists of numbers that are ordered solely on the basis of chance; it is used for drawing a random sample.

Simple Random Sampling

Simple random sampling requires some procedure that generates numbers or otherwise identifies cases strictly on the basis of chance. As you know, flipping a coin or rolling a die can be used to identify cases strictly on the basis of chance, but these procedures are not very efficient tools for drawing a sample. A **random number table**, such as the one in Appendix E, simplifies the process considerably. The

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:27 PM Page 161

researcher numbers all the elements in the sampling frame and then uses a systematic procedure for picking corresponding numbers from the random number table. (Practice Exercise 1 at the end of this chapter explains the process step-by-step.) Alternatively, a researcher may use a lottery procedure. Each case number is written on a small card, and then the cards are mixed up and the sample is selected from the cards.

When a large sample must be generated, these procedures are very cumbersome. Fortunately, a computer program can easily generate a random sample of any size. The researcher must first number all the elements to be sampled (the sampling frame) and then run the computer program to generate a random selection of the numbers within the desired range. The elements represented by these numbers are the sample.

Organizations that conduct phone surveys often draw random samples using another automated procedure, called **random digit dialing.** A machine dials random numbers within the phone prefixes corresponding to the area in which the survey is to be conducted. Random digit dialing is particularly useful when a sampling frame is not available. The researcher simply replaces any inappropriate number (e.g., those that are no longer in service or that are for businesses) with the next randomly generated phone number.

Random digit dialing The random dialing by a machine of numbers within designated phone prefixes, which creates a random sample for phone surveys.

The probability of selection in a true simple random sample is equal for each element. If a sample of 500 is selected from a population of 17,000 (i.e., a sampling frame of 17,000), then the probability of selection for each element is 500/17,000, or .03. Every element has an equal chance of being selected, just like the odds in a toss of a coin (1/2) or a roll of a die (1/6). Thus, simple random sampling is an "equal probability of selection method," or EPSEM.

Replacement sampling A method of sampling in which sample elements are returned to the sampling frame after being selected, so they may be sampled again. Random samples may be selected with or without replacement.

Simple random sampling can be done either with or without replacement sampling. In **replacement sampling**, each element is returned to the sampling frame after it is selected so that it may be sampled again. In sampling without replacement, each element selected for the sample is then excluded from the sampling frame. In practice, it makes no difference whether sampled elements are replaced after selection as long as the population is large and the sample is to contain only a small fraction of the population. Random sampling with replacement is, in fact, rarely used.

In a study involving simple random sampling, Bruce Link and his associates (1996) used random digit dialing to contact adult household members in the continental United States for an investigation of public attitudes and beliefs about homeless people. Sixty-three percent of the potential interviewees responded. The sample actually obtained was not exactly comparable with the population sampled: Compared with U.S. Census figures, the sample overrepresented women, people aged 25 to 54, married people, and those with more than a high school education; it underrepresented Latinos.

How does this sample strike you? Let's assess sample quality using the questions posed earlier in the chapter:

- *From what population were the cases selected?* There is a clearly defined population: the adult residents of the continental United States (who live in households with phones).

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:27 PM Page 162

- *What method was used to select cases from this population?* The case selection method is a random selection procedure, and there are no systematic biases in the sampling.
- *Do the cases that were studied represent, in the aggregate, the population from which they were selected?* The findings will very likely represent the population sampled because there were no biases in the sampling and a very large number of cases were selected. However, 37% of those selected for interviews could not be contacted or chose not to respond. This rate of nonresponse seems to create a small bias in the sample for several characteristics.

We must also consider the issue of cross-population generalizability: Do findings from this sample have implications for any larger group beyond the population from which the sample was selected? Because a representative sample of the entire U.S. adult population was drawn, this question has to do with cross-national generalizations. Link and his colleagues don't make any such generalizations. There's no telling what might occur in other countries with different histories of homelessness and different social policies.

Systematic random sampling A method of sampling in which sample elements are selected from a list or from sequential files.

Systematic Random Sampling

Systematic random sampling is a variant of simple random sampling. The first element is selected randomly from a list or from sequential files, and then every *n*th element is selected. This is a convenient method for drawing a random sample when the population elements are arranged sequentially. It is particularly efficient when the elements are not actually printed (i.e., there is no sampling frame) but instead are represented by folders in filing cabinets.

Systematic random sampling requires the following three steps:

1. The total number of cases in the population is divided by the number of cases required for the sample. This division yields the sampling interval, the number of cases from one sampled case to another. If 50 cases are to be selected out of 1,000, the **sampling interval** is 20; every 20th case is selected.
2. A number from 1 to 20 (or whatever the sampling interval is) is selected randomly. This number identifies the first case to be sampled, counting from the first case on the list or in the files.
3. After the first case is selected, every *n*th case is selected for the sample, where *n* is the sampling interval. If the sampling interval is not a whole number, the size of the sampling interval is varied systematically to yield the proper number of cases for the sample. For example, if the sampling interval is 30.5, the sampling interval alternates between 30 and 31. In almost all sampling situations, systematic random sampling yields what is essentially a simple random sample. The exception is a situation in which the sequence of elements is affected by **periodicity**—that is, the sequence varies in some regular, periodic pattern. For example, the houses in a new development with the same number of

Sampling interval The number of cases from one sampled case to another in a systematic random sample.

Periodicity A sequence of elements (in a list to be sampled) that varies in some regular, periodic pattern.

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

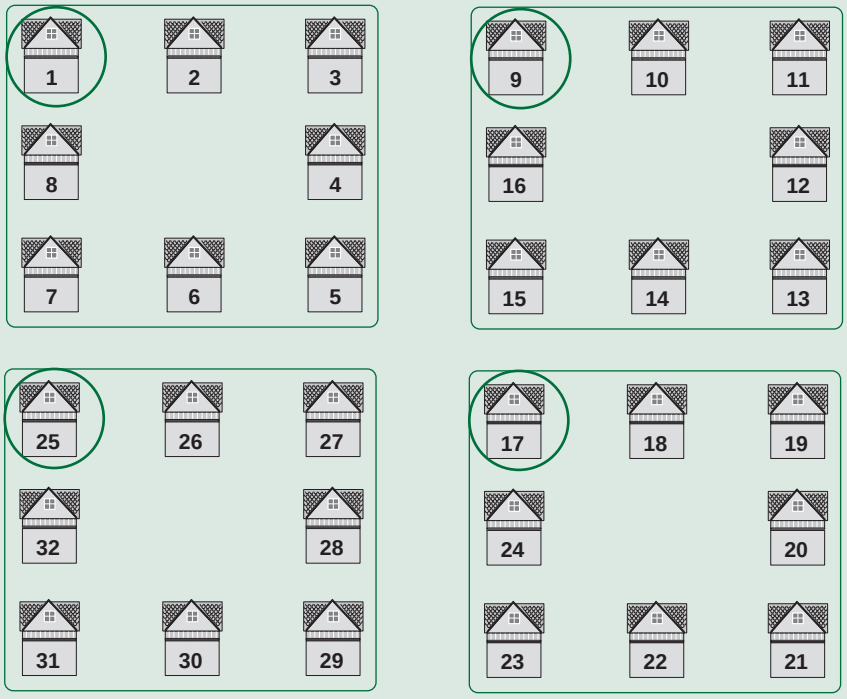
05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:27 PM Page 163

houses in each block (e.g., eight) may be listed by block, starting with the house in the northwest corner of each block and continuing clockwise. If the sampling interval is 8, the same as the periodic pattern, all the cases selected will be in the same position (see Exhibit 5.4). But in reality, periodicity and the sampling interval are rarely the same.

Stratified Random Sampling

Although all probability sampling methods use random sampling, some add steps to the sampling process to make sampling more efficient or easier. **Stratified random sampling** uses information known about the total population prior to sampling to make the sampling process more efficient. First, all elements in the population (i.e., in the sampling frame) are distinguished according to their value on some relevant characteristic. That characteristic forms the sampling strata. Next, elements are sampled randomly from within these strata. For example, race may be the basis for distinguishing individuals in some population of interest.

EXHIBIT 5.4 The Effect of Periodicity on Systematic Random Sampling



If the sampling interval is 8 for a study in this neighborhood, every element of the sample will be a house on the northwest corner—and thus the sample will be biased.

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:27 PM Page 164

Stratified random sampling A method of sampling in which sample elements are selected separately from population strata that are identified in advance by the researcher.

Within each racial category, individuals are then sampled randomly. Of course, using this method requires more information prior to sampling than is the case with simple random sampling. It must be possible to categorize each element in one and only one stratum, and the size of each stratum in the population must be known.

Proportionate stratified sampling Sampling method in which elements are selected from strata in exact proportion to their representation in the population.

This method is more efficient than drawing a simple random sample because it ensures appropriate representation of elements across strata. Imagine that you plan to draw a sample of 500 from an ethnically diverse neighborhood. The neighborhood population is 15% black, 10% Hispanic, 5% Asian, and 70% white. If you drew a simple random sample, you might end up with somewhat disproportionate numbers of each group. But if you created sampling strata based on race and ethnicity, you could randomly select cases from each stratum: 75 blacks (15% of the sample), 50 Hispanics (10%), 25 Asians (5%), and 350 whites (70%). By using **proportionate stratified sampling**, you would eliminate any possibility of sampling error in the sample's distribution of ethnicity. Each stratum would be represented exactly in proportion to its size in the population from which the sample was drawn (see Exhibit 5.5).

This is the strategy used by Brenda Booth et al. (2002) in a study of homeless adults in two Los Angeles county sites with large homeless populations. Specifically, Booth et al. (2002:432) selected subjects at random from homeless shelters, meal facilities, and from literally homeless populations on the streets. Respondents were sampled proportionately to their numbers in the downtown and Westside areas, as determined by a one-night enumeration. They were also sampled proportionately to their distribution across three nested sampling strata: the population using shelter beds, the population using meal facilities, and the unsheltered population using neither.

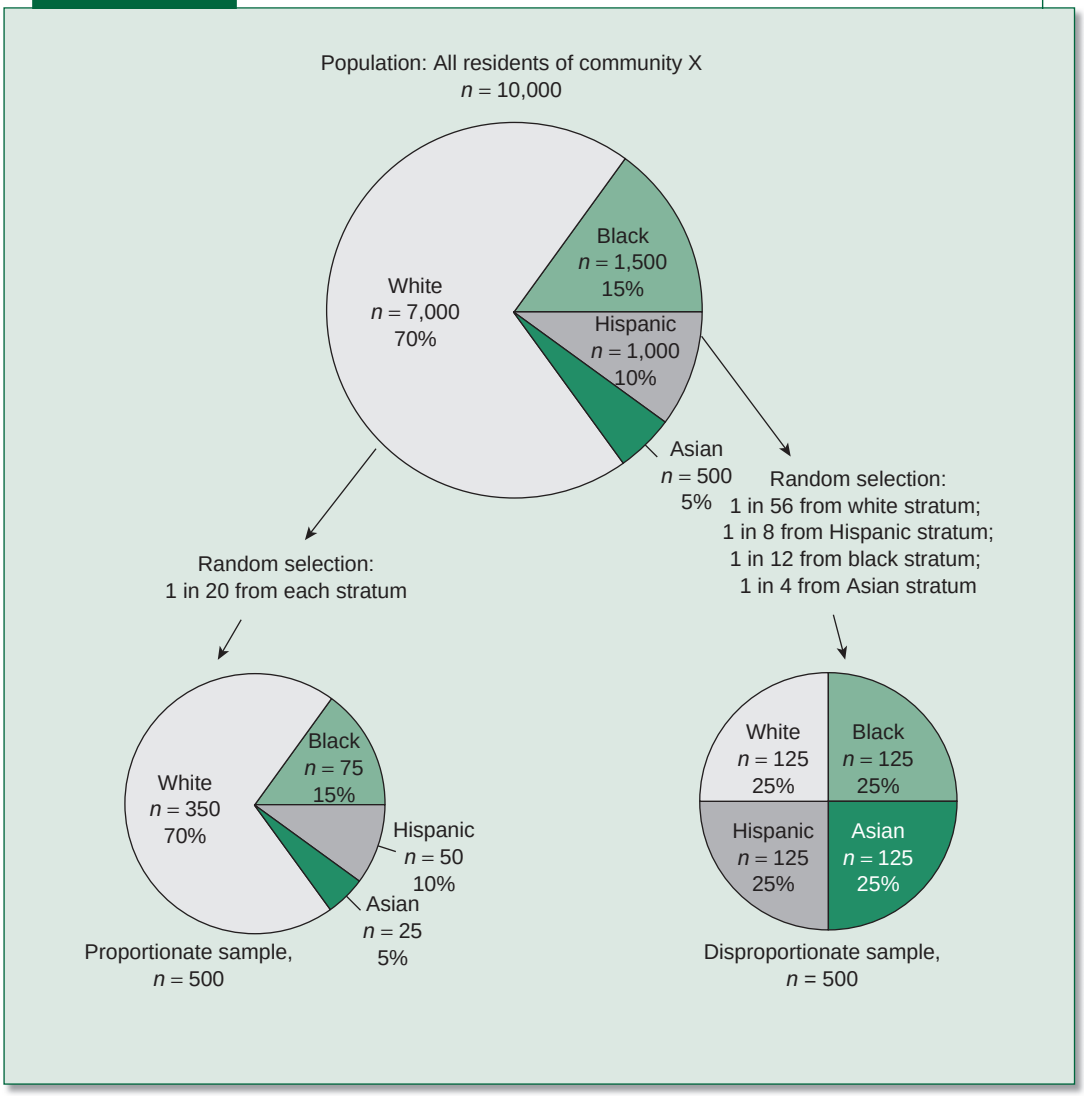
In **disproportionate stratified sampling**, the proportion of each stratum that is included in the sample is intentionally varied from what it is in the population. In the case of the sample stratified by ethnicity, you might select equal numbers of cases from each racial or ethnic group: 125 blacks (25% of the sample), 125 Hispanics (25%), 125 Asians (25%), and 125 whites (25%). In this type of sample, the probability of selection of every case is known but unequal between strata. You know what the proportions are in the population, and so you can easily adjust your combined sample statistics to reflect these true proportions. For instance, if you want to combine the ethnic groups and estimate the average income of the total population, you would have to "weight" each case in the sample. The weight is a number you multiply by the value of each case based on the stratum it is in. For example, you would multiply the incomes of all blacks in the sample by 0.6 (75/125), the incomes of all Hispanics by 0.4 (50/125), and so on. Weighting in this way reduces the influence of the oversampled strata and increases the influence of the undersampled strata to what they would have been if pure probability sampling had been used.

Disproportionate stratified sampling Sampling in which elements are selected from strata in different proportions from those that appear in the population.

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:28 PM Page 165

EXHIBIT 5.5 Stratified Random Sampling



Booth et al. (2002:432) included one element of *dis*proportionate random sampling in their otherwise proportionate random sampling strategy for homeless persons in Minneapolis: They oversampled women so that they comprised 26% of the sample, as compared with their actual percentage of 16% in the homeless population. Why would anyone select a sample that is so unrepresentative in the first place? The most common reason is to ensure that cases from smaller strata are included in the sample in sufficient numbers so as to allow separate statistical estimates and to facilitate comparisons between strata. Remember that one of the determinants

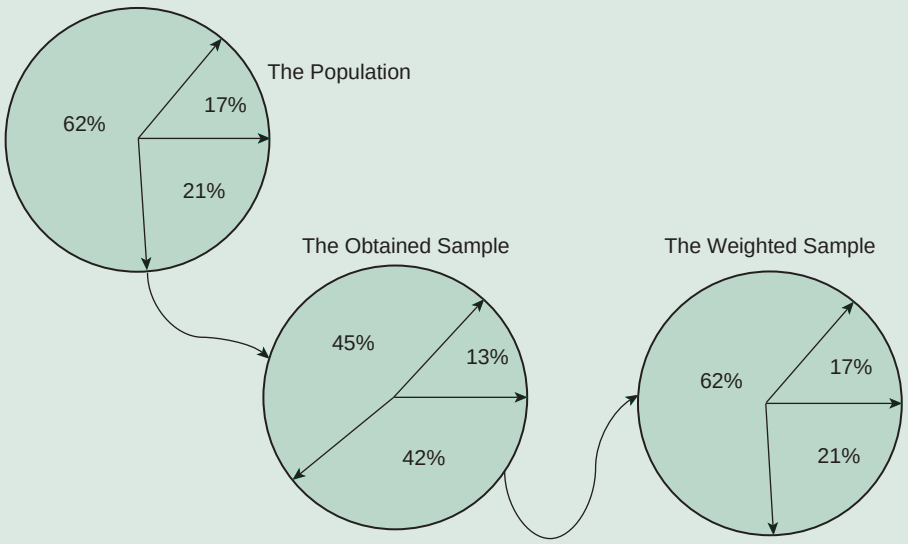
Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:28 PM Page 166

of sample quality is sample size. The same is true for subgroups within samples. If a key concern in a research project is to describe and compare the incomes of people from different racial and ethnic groups, then it is important that the researchers base the mean income of each group on enough cases to be a valid representation. If few members of a particular minority group are in the population, they need to be oversampled. Such disproportionate sampling may also result in a more efficient sampling design if the costs of data collection differ markedly between the strata or if the variability (heterogeneity) of the strata differs.

Weighting is also sometimes used to reduce the lack of representativeness of a sample that occurs due to nonresponse. On finding that the obtained sample does not represent the population in terms of some known characteristics such as, perhaps, gender or education, the researcher weights the cases in the sample so that the sample has the same proportions of men and women, or high school graduates and college graduates, as the complete population (see Exhibit 5.6). Keep in mind, though, that this procedure does not solve the problems caused by an unrepresentative sample because you still don't know what the sample composition should have been in terms of the other variables in your study; all you have done is to reduce the sample's unrepresentativeness in terms of the variables used in weighting. This may, in turn, make it more likely that the sample is representative of the population in terms of other characteristics, but you don't really know.

EXHIBIT 5.6 Weighting an Obtained Sample to Match a Population Proportion



Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:28 PM Page 167

Cluster Sampling

Cluster sampling is useful when a sampling frame of elements is not available, as often is the case for large populations spread out across a wide geographic area or among many different organizations. A **cluster** is a naturally occurring, mixed aggregate of elements of the population, with each element appearing in one, and only one, cluster. Schools could serve as clusters for sampling students, blocks could serve as clusters for sampling city residents, counties could serve as clusters for sampling the general population, and businesses could serve as clusters for sampling employees.

Cluster sampling Sampling in which elements are selected in two or more stages, with the first stage being the random selection of naturally occurring clusters and the last stage being the random selection of elements within clusters.

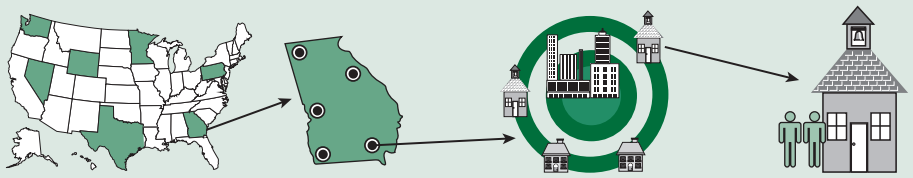
Cluster A naturally occurring, mixed aggregate of elements of the population.

Drawing a cluster sample is, at least, a two-stage procedure. First, the researcher draws a random sample of clusters. A list of clusters should be much easier to obtain than a list of all the individuals in each cluster in the population. Next, the researcher draws a random sample of elements within each selected cluster. Because only a fraction of the total clusters are involved, obtaining the sampling frame at this stage should be much easier.

In a cluster sample of city residents, for example, blocks could be the first-stage clusters. A research assistant could walk around each selected block and record the addresses of all occupied dwelling units. Or, in a cluster sample of students, a researcher could contact the schools selected in the first stage and make arrangements with the registrar to obtain lists of students at each school. Cluster samples often involve multiple stages (see Exhibit 5.7), with clusters within clusters, as when a national sample of individuals might involve first sampling states, then geographic units within those states, then dwellings within those units, and finally, individuals within the dwellings. In multistage cluster sampling, the clusters at the first stage of sampling are termed the *primary sampling units* (Levy & Lemeshow 1999:228).

How many clusters should be selected, and how many individuals within each cluster should be selected? As a general rule, the sample will be more similar to the entire population

EXHIBIT 5.7 Multistage Cluster Sampling



Stage 1: Randomly select states

Stage 2: Randomly select cities, towns, and counties within those states

Stage 3: Randomly select schools within those cities and towns

Stage 4: Randomly select students within each school

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:28 PM Page 168

if the researcher selects as many clusters as possible—even though this will mean the selection of fewer individuals within each cluster. Unfortunately, this strategy also maximizes the cost of the sample for studies using in-person interviews. The more clusters a researcher selects, the more time and money will have to be spent traveling to the different clusters to reach the individuals for interviews.

The calculation of how many clusters to sample and how many individuals are within the clusters is also affected by the degree of similarity of individuals within clusters: The more similar the individuals are within the clusters, the fewer the number of individuals needed to represent each cluster. So if you set out to draw a cluster sample, be sure to consider how similar individuals are within the clusters as well as how many clusters you can afford to include in your sample.

Cluster sampling is a very popular method among survey researchers, but it has one general drawback: Sampling error is greater in a cluster sample than in a simple random sample, because there are two steps involving random selection rather than just one. This sampling error increases as the number of clusters decreases, and it decreases as the homogeneity of cases per cluster increases. In sum, it's better to include as many clusters as possible in a sample, and it's more likely that a cluster sample will be representative of the population if cases are relatively similar within clusters.

Probability Sampling Methods Compared

Can you now see why researchers often prefer to draw a stratified random sample or a cluster sample rather than a simple random sample? Exhibit 5.8 should help you remember the key features of these different types of sample and to determine when each is most appropriate.

Many professionally designed surveys use combinations of clusters and stratified probability sampling methods. For example, Peter Rossi (1989) drew a disproportionate stratified cluster sample of shelter users for a homelessness study in Chicago (see Exhibit 5.9). The shelter sample was stratified by size, with smaller shelters having a smaller likelihood of selection than larger shelters. In fact, the larger shelters were all selected; they had a probability of selection of 1.0. Within the selected shelters, shelter users were then sampled using a systematic random selection procedure (except in the small shelters, in which all persons were interviewed). Homeless persons living on the streets were also sampled randomly. In the first stage, city blocks were classified into strata based on the likely concentration of homeless persons (estimated by several knowledgeable groups). Blocks were then picked randomly within these

EXHIBIT 5.8 Features of Probability Sampling Methods

Feature	*Simple*	*Systematic*	*Stratified*	*Cluster*
Unbiased selection of cases	Yes	Yes	Yes	Yes
Sampling frame required	Yes	No	Yes	No
Ensures representation of key strata	No	No	Yes	No
Uses natural grouping of cases	No	No	No	Yes
Reduces sampling costs	No	No	No	Yes
Sampling error compared with SRS	–	Same	Lower	Higher

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:28 PM Page 169

strata and, on the survey night between 1 a.m. and 6 a.m., teams of interviewers screened each person found outside on that block for his or her homeless status. Persons identified as homeless were then interviewed (and given $5 for their time). The rate of response for two different samples (fall and winter) in the shelters and on the streets was between 73% and 83%.

How would we evaluate the Chicago homeless sample, using the sample evaluation questions?

- *From what population were the cases selected?* The population was clearly defined for each cluster.
- *What method was used to select cases from this population?* The random selection method was carefully described.
- *Do the cases that were studied represent, in the aggregate, the population from which they were selected?* The unbiased selection procedures make us reasonably confident in the representativeness of the sample, although we know little about the nonrespondents and therefore may justifiably worry that some types of homeless persons were missed.

Cross-population generalization seems to be reasonable with this sample, because it seems likely that the findings reflect general processes involving homeless persons. Rossi (1989) clearly thought so, because his book's title refers to homelessness in America, not just in Chicago.

Nonprobability Sampling Methods

Nonprobability sampling methods are often used in qualitative research; they also are used in quantitative studies when researchers are unable to use probability selection methods. In qualitative research, a focus on one setting or a very small sample allows a more intensive portrait of activities and actors, but it also limits field researchers' ability to generalize and

EXHIBIT 5.9 Chicago Shelter Universe and Shelter Samples, Fall and Winter Surveys

A. Shelter Universe and Samples

	Fall	*Winter*
Eligible shelters in universe	28	45
Universe bed capacities	1,573	2,001
Shelters drawn in sample	22	27

B. Details of Winter Shelter Sample

Shelter Size Classification	*Number in Universe*	*Number in Sample*	*Occupant Sampling Ratio*
Large (37 or more beds)	17	17	0.25
Medium (18–33 beds)	12	6	0.50
Small (under 18 beds)	16	4	1.00

Note: Shelters were drawn with probabilities proportionate to size, with residents sampled disproportionately within shelters to form a self-weighting sample. Sampling ratios for the Phase 2 sample are given in Panel B.

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:28 PM Page 170

lowers the confidence that others can place in these generalizations. The use of nonprobability sampling methods in quantitative research too often reflects a lack of concern with generalizability or a lack of understanding of the importance of probability-based sampling.

There are four common nonprobability sampling methods: availability sampling, quota sampling, purposive sampling, and snowball sampling. Because these methods do not use a random selection procedure, we cannot expect a sample selected with any of these methods to yield a representative sample. They should not be used in quantitative studies if a probability-based method is feasible. Nonetheless, these methods are useful when random sampling is not possible, when a research question calls for an intensive investigation of a small population, or when a researcher is performing a preliminary, exploratory study.

Availability Sampling

Elements are selected for **availability sampling** because they're available or easy to find. Thus, this sampling method is also known as a haphazard, accidental, or convenience sample.

Availability sampling Sampling in which elements are selected on the basis of convenience.

There are many ways to select elements for an availability sample: Standing on street corners and talking to whoever walks by, asking questions of employees who have time to talk when they pick up their paycheck at a personnel office, or approaching particular individuals at opportune times while observing activities in a social setting. You may find yourself interviewing available students at campus hangouts as part of a course assignment. To study sexual risk-taking among homeless youth in Minneapolis, Linda Halcón and Alan Lifson (2004:73) hired very experienced street youth outreach workers who approached youth known or suspected to be homeless and asked if they would be willing to take part in a 20- to 30-minute interview.

The interviewers then conducted the 44-question interview, after which they gave respondents some risk reduction and referral information and a $20 voucher.

A participant observation study of a group may require no more sophisticated approach. When Philippe Bourgois, Mark Lettiere, and James Quesada (1997) studied homeless heroin addicts in San Francisco, they immersed themselves in a community of addicts living in a public park. These addicts became the availability sample.

An availability sample is often appropriate in social research—for example, when a field researcher is exploring a new setting and trying to get some sense of the prevailing attitudes or when a survey researcher conducts a preliminary test of a new set of questions.

Now I'd like you to use the sample evaluation questions to evaluate person-in-the-street interviews of the homeless. If your answers are something like "The population was unknown," "The method for selecting cases was haphazard," and "The cases studied do not represent the population," you're right! There is no clearly definable population from which the respondents were drawn, and no systematic technique was used to select the respondents. There certainly is not much likelihood that the interviewees represent the distribution of sentiment among homeless persons in the Boston area or of welfare mothers or of impoverished rural migrants or of whatever we imagine the relevant population is.

In a similar vein, perhaps person-in-the street comments to news reporters suggest something about what homeless persons think, or maybe they don't; we can't really be sure. But let's give reporters their due: If they just want to have a few quotes to make their story more

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:28 PM Page 171

appealing, nothing is wrong with their sampling method. However, their approach gives us no basis for thinking that we have an overview of the community sentiment. The people who happen to be available in any situation are unlikely to be just like those who are unavailable. We can't be at all certain that what we learn can be generalized with any confidence to a larger population of concern.

Availability sampling often masquerades as a more rigorous form of research. Popular magazines periodically survey their readers by printing a questionnaire for readers to fill out and mail in. A follow-up article then appears in the magazine under a title such as "What You Think About Intimacy in Marriage." If the magazine's circulation is large, a large sample can be achieved in this way. The problem is that usually only a tiny fraction of readers return the questionnaire, and these respondents are probably unlike other readers who did not have the interest or time to participate. So the survey is based on an availability sample. Even though the follow-up article may be interesting, we have no basis for thinking that the results describe the readership as a whole—much less the population at large.

Do you see now why availability sampling differs so much from random sampling methods, which require that "nothing but chance" affects the actual selection of cases? What makes availability sampling "haphazard" is precisely that a great many things other than chance can affect the selection of cases, ranging from the prejudices of the research staff to the work schedules of potential respondents. To truly leave the selection of cases up to chance, we have to design the selection process very carefully so that other factors are not influential. There's nothing "haphazard" about selecting cases randomly.

Quota Sampling

Quota sampling is intended to overcome the most obvious flaw of availability sampling—that the sample will just consist of whoever or whatever is available, without any concern for its similarity to the population of interest. The distinguishing feature of a quota sample is that quotas are set to ensure that the sample represents certain characteristics in proportion to their prevalence in the population.

Quota sampling A nonprobability sampling method in which elements are selected to ensure that the sample represents certain characteristics in proportion to their prevalence in the population.

Suppose that you wish to sample adult residents of a town in a study of support for a tax increase to improve the town's schools. You know from the town's annual report what the proportions of town residents are in terms of gender, race, age, and number of children. You think that each of these characteristics might influence support for new school taxes, so you want to be sure that the sample includes men, women, whites, blacks, Hispanics, Asians, older people, younger people, big families, small families, and childless families in proportion to their numbers in the town population.

This is where quotas come in. Let's say that 48% of the town's adult residents are men and 52% are women, and that 60% are employed, 5% are unemployed, and 35% are out of the labor force. These percentages and the percentages corresponding to the other characteristics become the quotas for the sample. If you plan to include a total of 500 residents in your sample, 240 must be men (48% of 500), 260 must be women, 300 must be employed, and so on. You may even set more refined quotas, such as certain numbers of employed women, employed men, unemployed men, and so on. With the quota list in hand, you (or your research

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:29 PM Page 172

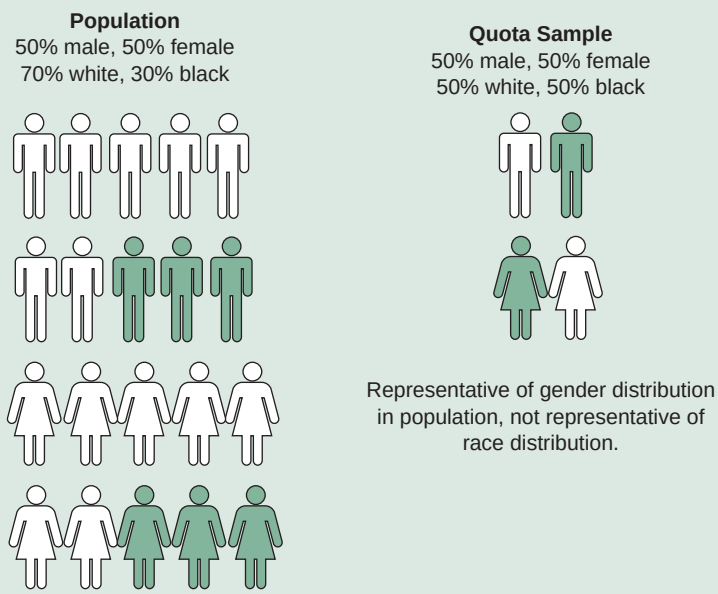
staff) can now go out into the community looking for the right number of people in each quota category. You may go door to door, bar to bar, or just stand on a street corner until you have surveyed 240 men, 260 women, and so on.

The problem is that even when we know that a quota sample is representative of the particular characteristics for which quotas have been set, we have no way of knowing if the sample is representative in terms of any other characteristics. In Exhibit 5.10, for example, quotas have been set for gender only. Under these circumstances, it's no surprise that the sample is representative of the population only in terms of gender, not in terms of race. Interviewers are only human; they may avoid potential respondents with menacing dogs in the front yard, or they could seek out respondents who are physically attractive or who look like they'd be easy to interview. Realistically, researchers can set quotas for only a small fraction of the characteristics relevant to a study, so a quota sample is really not much better than an availability sample (although following careful, consistent procedures for selecting cases within the quota limits always helps).

This last point leads me to another limitation of quota sampling: You must know the characteristics of the entire population to set the right quotas. In most cases, researchers know what the population looks like in terms of no more than a few of the characteristics relevant to their concerns—and in some cases, they have no such information on the entire population.

If you're now feeling skeptical of quota sampling, you've gotten the drift of my remarks. Nonetheless, in some situations, establishing quotas can add rigor to sampling procedures. It's almost always better to maximize possibilities for comparison in research, and quota sampling techniques can help qualitative researchers do this. For instance, Doug Timmer, Stanley Eitzen, and Kathryn Talley (1993:7) interviewed homeless persons in several cities and other locations

EXHIBIT 5.10 Quota Sampling



Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:29 PM Page 173

for their book on the sources of homelessness. Persons who were available were interviewed, but the researchers paid some attention to generating a diverse sample. They interviewed 20 homeless men who lived on the streets without shelter and 20 mothers who were found in family shelters. About half of those the researchers selected in the street sample were black, and about half were white. Although the researchers did not use quotas to try to match the distribution of characteristics among the total homeless population, their informal quotas helped ensure some diversity in key characteristics.

Does quota sampling remind you of stratified sampling? It's easy to understand why, since they both select sample members based, in part, on the basis of one or more key characteristics. Exhibit 5.11 summarizes the differences between quota sampling and stratified random sampling. The key difference, of course, is quota sampling's lack of random selection.

Purposive Sampling

In **purposive sampling**, each sample element is selected for a purpose, usually because of the unique position of the sample elements. Purposive sampling may involve studying the entire population of some limited group (directors of shelters for homeless adults) or a subset of a population (mid-level managers with a reputation for efficiency). Or a purposive sample may be a "key informant survey," which targets individuals who are particularly knowledgeable about the issues under investigation.

Purposive sampling A nonprobability sampling method in which elements are selected for a purpose, usually because of their unique position.

Herbert Rubin and Irene Rubin (1995) suggest three guidelines for selecting informants when designing any purposive sampling strategy. Informants should be

- "knowledgeable about the cultural arena or situation or experience being studied,"
- "willing to talk," and
- "represent[ative of] the range of points of view." (p. 66)

In addition, Rubin and Rubin (1995) suggest continuing to select interviewees until you can pass two tests:

- *Completeness:* "What you hear provides an overall sense of the meaning of a concept, theme, or process." (p. 72)
- *Saturation:* "You gain confidence that you are learning little that is new from subsequent interview[s]." (p. 73)

EXHIBIT 5.11 Comparison of Stratified and Quota Sampling Methods

Feature	*Stratified*	*Quota*
Unbiased (random) selection of cases	Yes	No
Sampling frame required	Yes	No
Ensures representation of key strata	Yes	Yes

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:29 PM Page 174

Adhering to these guidelines will help ensure that a purposive sample adequately represents the setting or issues studied.

Of course, purposive sampling does not produce a sample that represents some larger population, but it can be exactly what is needed in a case study of an organization, community, or some other clearly defined and relatively limited group. In an intensive organizational case study, a purposive sample of organizational leaders might be complemented with a probability sample of organizational members. Before designing her probability samples of hospital patients and homeless persons, Dee Roth (1990:146–147) interviewed a purposive sample of 164 key informants from organizations that had contact with homeless people in each of the counties she studied.

Snowball Sampling

Snowball sampling is useful for hard-to-reach or hard-to-identify populations for which there is no sampling frame, but the members of which are somewhat interconnected (at least some members of the population know each other). It can be used to sample members of groups such as drug dealers, prostitutes, practicing criminals, participants in Alcoholics Anonymous groups, gang leaders, informal organizational leaders, and homeless persons. It may also be used for charting the relationships among members of some group (a sociometric study), for exploring the population of interest prior to developing a formal sampling plan, and for developing what becomes a census of informal leaders of small organizations or communities. However, researchers using snowball sampling normally cannot be confident that their sample represents the total population of interest, so generalizations must be tentative.

Snowball sampling A method of sampling in which sample elements are selected as they are identified by successive informants or interviewees.

Rob Rosenthal (1994) used snowball sampling to study homeless persons living in Santa Barbara, California:

> I began this process by attending a meeting of homeless people I had heard about through my housing advocate contacts. . . . One homeless woman . . . invited me to . . . where she promised to introduce me around. Thus a process of snowballing began. I gained entree to a group through people I knew, came to know others, and through them gained entree to new circles. (pp. 178, 180)

One problem with this technique is that the initial contacts may shape the entire sample and foreclose access to some members of the population of interest:

> Sat around with [my contact] at the Tree. Other people come by, are friendly, but some regulars, especially the tougher men, don't sit with her. Am I making a mistake by tying myself too closely to her? She lectures them a lot. (Rosenthal 1994:181)

More systematic versions of snowball sampling can reduce the potential for bias. For example, "respondent-driven sampling" gives financial incentives to respondents to recruit peers (Heckathorn 1997). Limitations on the number of incentives that any one respondent can receive increase the sample's diversity. Targeted incentives can steer the sample to include

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:29 PM Page 175

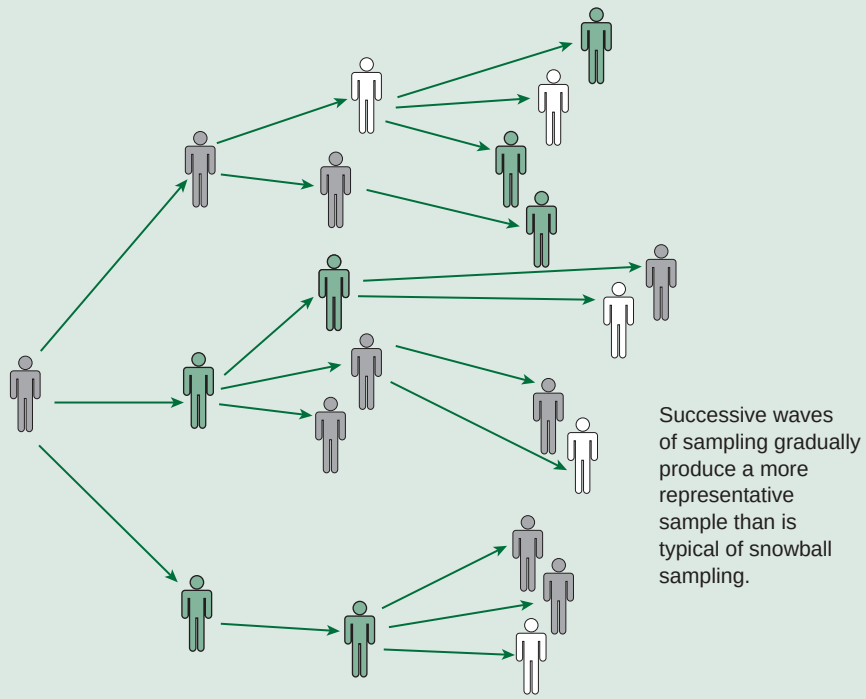
specific subgroups. When the sampling is repeated through several waves, with new respondents bringing in more peers, the composition of the sample converges on a more representative mix of characteristics than would occur with uncontrolled snowball sampling. Exhibit 5.12 shows how the sample spreads out through successive recruitment waves to an increasingly diverse pool (Heckathorn 1997:178). Exhibit 5.13 shows that even if the starting point were all white persons, respondent-driven sampling would result in an appropriate ethnic mix from an ethnically diverse population (Heckathorn 2002:17).

Lessons About Sample Quality

Some lessons are implicit in my evaluations of the samples in this chapter:

- We can't evaluate the quality of a sample if we don't know what population it is supposed to represent. If the population is unspecified because the researchers were never clear about the population they were trying to sample, then we can safely conclude that the sample itself is no good.

EXHIBIT 5.12 Respondent-Driven Sampling

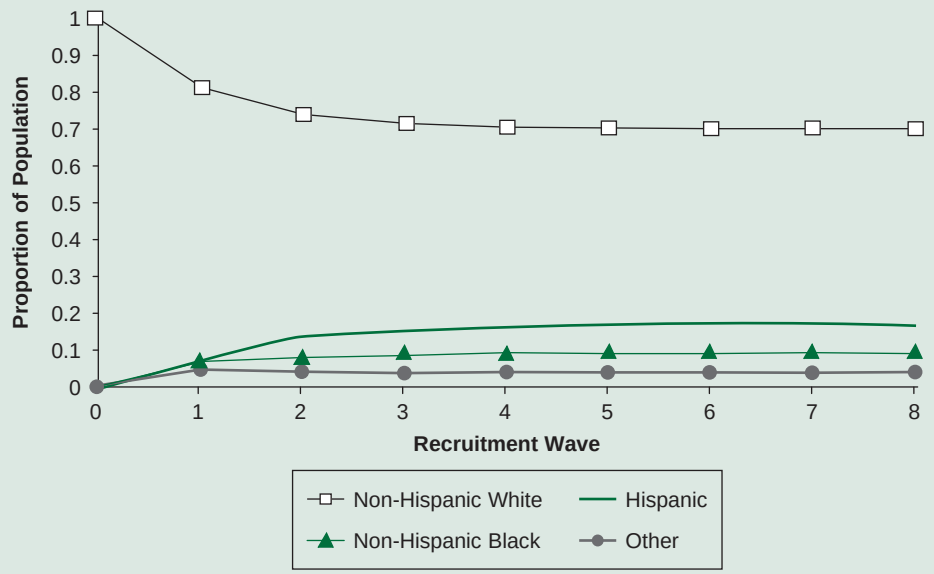


Instructions to respondents:
"We'll pay you $5 each for up to three names, but only one of those names can be somebody from your own town. The others have to be from somewhere else."

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:29 PM Page 176

EXHIBIT 5.13 Convergence of Respondent-Driven Sample to True Ethnic Proportions in Population, After Starting With Only Whites



- We can't evaluate the quality of a sample if we don't know how cases in the sample were selected from the population. If the method was specified, we then need to know whether cases were selected in a systematic fashion and on the basis of chance. In any case, we know that a haphazard method of sampling (as in person-on-the-street interviews) undermine generalizability.
- Sample quality is determined by the sample actually obtained, not just by the sampling method itself. If many of the people selected for our sample are nonrespondents or people (or other entities) who do not participate in the study although they have been selected for the sample, the quality of our sample is undermined—even if we chose the sample in the best possible way.
- We need to be aware that even researchers who obtain very good samples may talk about the implications of their findings for some group that is larger than, or just different from, the population they actually sampled. For example, findings from a representative sample of students in one university often are discussed as if they tell us about university students in general. And maybe they do; we just don't know for sure.
- A sample that allows for comparisons involving theoretically important variables is better than one that does not allow such comparisons. Even when we study people or social processes in depth, it is best to select individuals or settings with an eye to how useful they will be for examining relationships. Limiting an investigation to just one setting or just one type of person will inevitably leave us wondering what it is that makes a difference.

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:29 PM Page 177

Generalizability in Qualitative Research

Qualitative research often focuses on populations that are hard to locate or very limited in size. In consequence, such nonprobability sampling methods as availability sampling and snowball sampling are often used. Janet Wards Schofield (2002) suggests ways of increasing the generalizability of the samples obtained in such situations:

> *Studying the Typical.* Choosing sites on the basis of their fit with a typical situation is far preferable to choosing on the basis of convenience. (p. 181)
>
> *Performing Multisite Studies.* A finding emerging repeatedly in the study of numerous sites would appear to be more likely to be a good working hypothesis about some as yet unstudied site than a finding emerging from just one or two sites. . . . Generally speaking, a finding emerging from the study of several very heterogenous sites would be more . . . likely to be useful in understanding various other sites than one emerging from the study of several very similar sites. (p. 184)

The effort of some qualitative researchers to understand the particulars of a situation in depth, as an important object of inquiry in itself, also leads some to question the value of generalizability, as most researchers understand it. In the words of sociologist Norman Denzin,

> The interpretivist rejects generalization as a goal and never aims to draw randomly selected samples of human experience. . . . Every instance of social interaction . . . represents a slice from the life world that is the proper subject matter for interpretive inquiry. (Denzin cited in Schofield 2002:173)

SAMPLING DISTRIBUTIONS

A well-designed probability sample is one that is likely to be representative of the population from which it was selected. But as you've seen, random samples still are subject to sampling error owing just to chance. To deal with that problem, social researchers take into account the properties of a sampling distribution, a hypothetical distribution of a statistic across all the random samples that could be drawn from a population. Any single random sample can be thought of as just one of an infinite number of random samples that, in theory, could have been selected from the population. If we had the finances of Gatsby and the patience of Job and were able to draw an infinite number of samples, and we calculated the same type of statistic for each of these samples, we would then have a sampling distribution. Understanding sampling distributions is the foundation for understanding how statisticians can estimate sampling error.

What does a sampling distribution look like? Because a sampling distribution is based on some statistic calculated for different samples, we need to choose a statistic. Let's focus on the arithmetic average, or mean. I will explain the calculation of the mean in Chapter 14, but you may already be familiar with it: You add up the values of all the cases and divide by the total number of cases. Let's say you draw a random sample of 500 families and find that their average (mean) family income is $58,239. Imagine that you then draw another random sample. That sample's mean family income might be $60,302. Imagine marking these two

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:29 PM Page 178

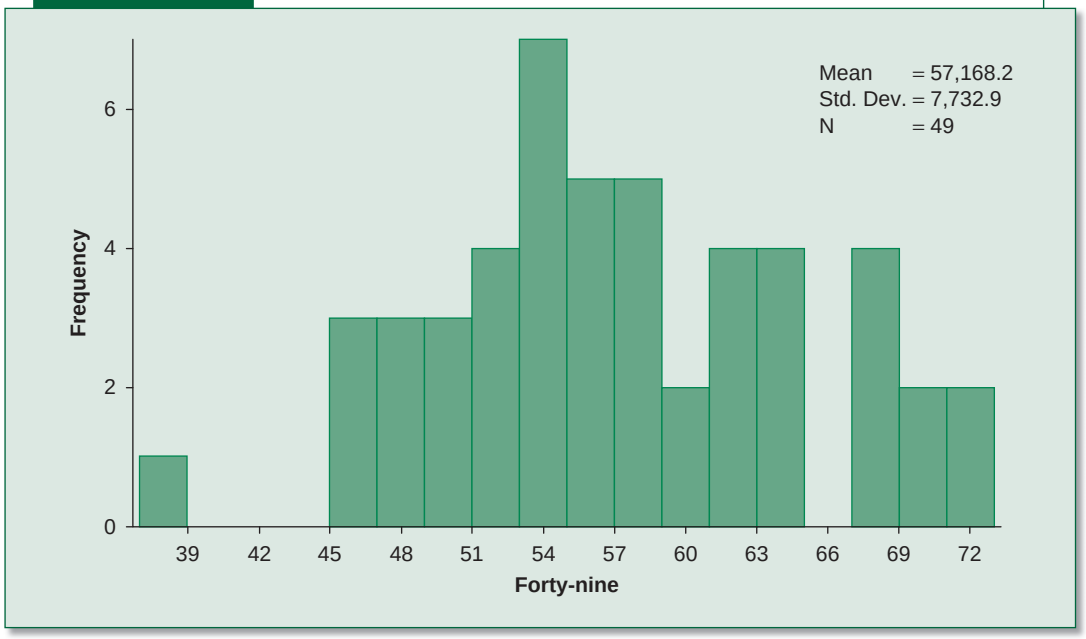
means on graph paper and then drawing more random samples and marking their means on the graph. The resulting graph would be a sampling distribution of the mean.

Exhibit 5.14 demonstrates what happened when I did something very similar to what I have just described—not with an infinite number of samples and not from a large population but through the same process using the 2006 General Social Survey (GSS) sample as if it were a population. First, I drew 49 different random samples, each consisting of 30 cases, from the 2006 GSS. (The standard notation for the number of cases in each sample is $n = 30$.) Then I calculated for each random sample the approximate mean family income (approximate because the GSS does not record actual income in dollars). I then graphed the means of the 49 samples. Each bar in Exhibit 5.13 shows how many samples had a particular family income. The mean for the population (the total GSS sample) is $59,213, and you can see that many of the samples in the sampling distribution are close to this value. However, although many of the sample means are close to the population mean, some are quite far from it. If you had calculated the mean from only one sample, it could have been anywhere in this sampling distribution, but it is unlikely to have been far from the population mean—that is, unlikely to have been close to either end (or "tail") of the distribution.

Estimating Sampling Error

We don't actually observe sampling distributions in real research; researchers just draw the best sample they can and then are stuck with the results—one sample, not a distribution of samples. A sampling distribution is a theoretical distribution. However, we can use the

EXHIBIT 5.14 Partial Sampling Distribution: Mean Family Income



Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:30 PM Page 179

properties of sampling distributions to calculate the amount of sampling error that was likely with the random sample used in a study. The tool for calculating sampling error is called **inferential statistics**.

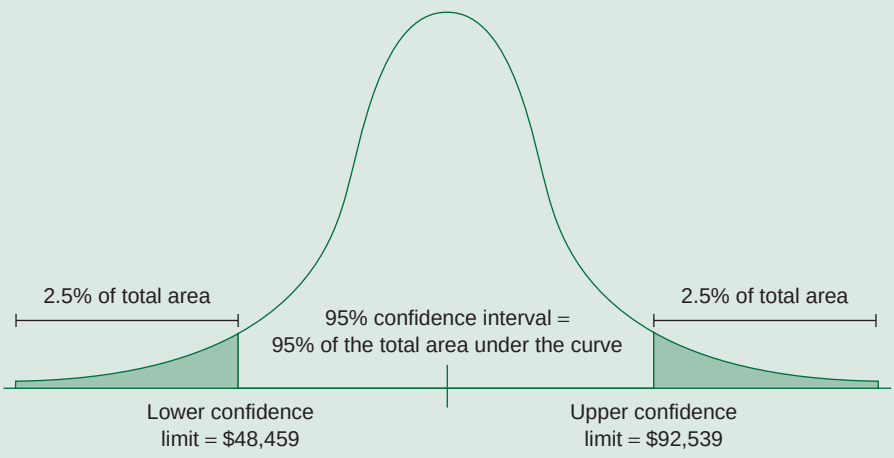
Sampling distributions for many statistics, including the mean, have a "normal" shape. A graph of a normal distribution looks like a bell, with one "hump" in the middle, centered on the population mean, and the number of cases tapering off on both sides of the mean. Note that a normal distribution is symmetric: If you folded it in half at its center (at the population mean), the two halves would match perfectly. This shape is produced by **random sampling error**—variation owing purely to chance. The value of the statistic varies from sample to sample because of chance, so higher and lower values are equally likely.

The partial sampling distribution in Exhibit 5.14 does not have a completely normal shape because it involves only a small number of samples (49), each of which has only 30 cases. Exhibit 5.15 shows what the sampling distribution of family incomes would look like if it formed a perfectly normal distribution—if, rather than 49 random samples, I had selected thousands of random samples.

Inferential statistics A mathematical tool for estimating how likely it is that a statistical result based on data from a random sample is representative of the population from which the sample is assumed to have been selected.

Random sampling error (chance sampling error) Differences between the population and the sample that are due only to chance factors (random error), not to systematic sampling error. Random sampling error may or may not result in an unrepresentative sample. The magnitude of sampling error due to chance factors can be estimated statistically.

EXHIBIT 5.15 Normal Sampling Distribution: Mean Family Income



Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:30 PM Page 180

Sample statistic The value of a statistic, such as a mean, computed from sample data.

Population parameter The value of a statistic, such as a mean, computed using the data for the entire population; a sample statistic is an estimate of a population parameter.

The properties of a sampling distribution facilitate the process of statistical inference. In the sampling distribution, the most frequent value of the **sample statistic**—the statistic (such as the mean) computed from sample data—is identical to the **population parameter**—the statistic computed for the entire population. In other words, we can have a lot of confidence that the value at the peak of the bell curve represents the norm for the entire population. A population parameter also may be termed the *true value* for the statistic in that population. A sample statistic is an estimate of a population parameter.

In a normal distribution, a predictable proportion of cases fall within certain ranges. Inferential statistics takes advantage of this feature and allows researchers to estimate how likely it is that, given a particular sample, the true population value will be within some range of the statistic. For example, a statistician might conclude from a sample of 30 families that *we can be 95% confident that the true mean family income in the total population is between $33,813 and $53,754.* The interval from $33,813 to $53,754 would then be called the "95% confidence interval for the mean." The lower ($33,813) and upper ($53,754) bounds of this interval are termed the *confidence limits.* Exhibit 5.15 marks such confidence limits, indicating the range that encompasses 95% of the area under the normal curve; 95% of all sample means would fall within this range, as does the mean of our hypothetical sample of 30 cases.

Although all normal distributions have these same basic features, they differ from one another in the extent to which they cluster around the mean. A sampling distribution is more compact when it is based on larger samples. Stated another way, we can be more confident in estimates based on larger random samples because we know that a larger sample creates a more compact sampling distribution. Compare the two sampling distributions of mean family income shown in Exhibit 5.16. Both depict the results for about 50 samples. However, in one study, each sample consisted of 100 families, and in the other study each sample consisted of only 5 families. Clearly, the larger samples result in a sampling distribution that is much more tightly clustered around the mean (range of 34 to 44) than is the case with the smaller samples (range of 17 to 57). The 95% confidence interval for mean family income for the entire 2006 GSS sample of 3,873 cases (the ones that had valid values of family income) was $57,416 to $61,009—an interval only $3,593 wide. But the 95% confidence interval for the mean family income in one GSS subsample of 100 cases was much wider, with limits of $48,891 and $72,501. And for a subsample of only 5 cases, the 95% confidence interval was very broad indeed: from $16,566 to $161,310. As you can see, such small samples result in statistics that actually give us very little useful information about the population.

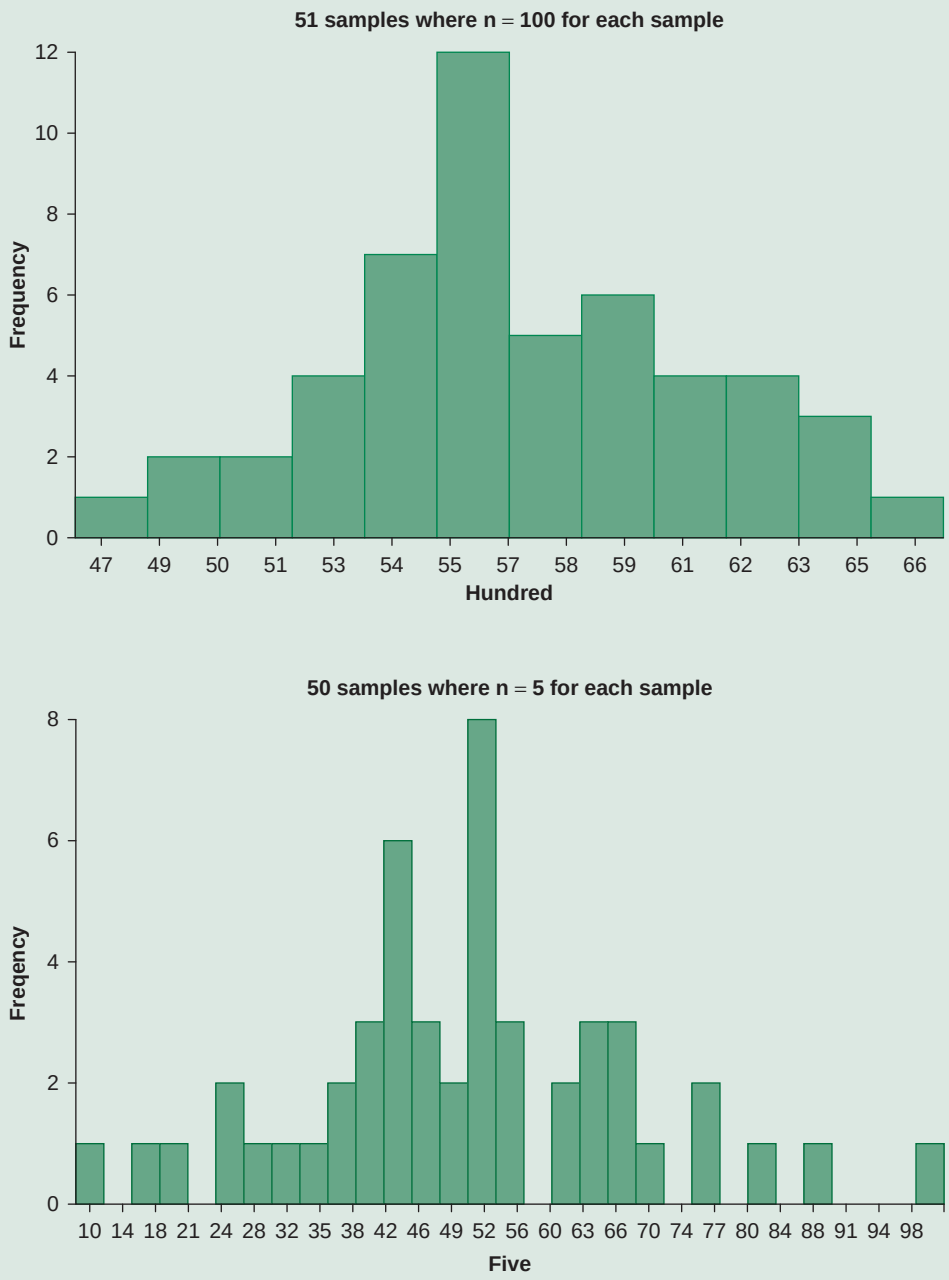
Other confidence intervals, such as the 99% confidence interval, can be reported. As a matter of convention, statisticians use only the 95%, 99%, and 99.9% confidence limits to estimate the range of values that are likely to contain the true value. These conventional limits reflect the conservatism inherent in classical statistical inference: Don't make an inferential statement unless you are very confident (at least 95% confident) that it is correct.

The less precise an estimate of a particular statistic from a particular sample is, the more confident we can be—and the wider the confidence interval. As I mentioned above, the 95%

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:30 PM Page 181

EXHIBIT 5.16 The Effect of Sample Size on Sampling Distribution



Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:30 PM Page 182

confidence interval for the entire 2006 GSS sample is $57,416 to $61,009 (a width of $3,593); the 99% confidence interval is $56,848 to $61,578 (a width of $4,730).

I will explain how to calculate confidence intervals in Chapter 14. You will find it easier to understand this procedure after you have learned some of the basic statistics that I have introduced in that chapter. If you have already completed a statistics course, you might want to turn now to Chapter 14's confidence interval section for a quick review. In any case, you should now have a sense of how researchers make inferences from a random sample of a population.

Determining Sample Size

You have learned now that more confidence can be placed in the generalizability of statistics from larger samples, so that you may be eager to work with random samples that are as large as possible. Unfortunately, researchers often cannot afford to sample a very large number of cases. Therefore, they try to determine during the design phase of their study how large a sample they must have to achieve their purposes. They have to consider the degree of confidence desired, the homogeneity of the population, the complexity of the analysis they plan, and the expected strength of the relationships they will measure.

- The less sampling error desired, the larger the sample size must be.
- Samples of more homogeneous populations can be smaller than samples of more diverse populations. Stratified sampling uses prior information on the population to create more homogeneous population strata from which the sample can be selected, so stratified samples can be smaller than simple random samples.
- If the only analysis planned for a survey sample is to describe the population in terms of a few variables, a smaller sample is required than if a more complex analysis involving sample subgroups is planned. If much of the analysis will focus on estimating the characteristics of subgroups within the sample, it is the size of the subgroups that must be considered, not the size of the total sample (Levy & Lemeshow 1999:74).
- When the researchers expect to find very strong relationships among the variables when they test hypotheses, they will need a smaller sample to detect these relationships than if they expect weaker relationships.

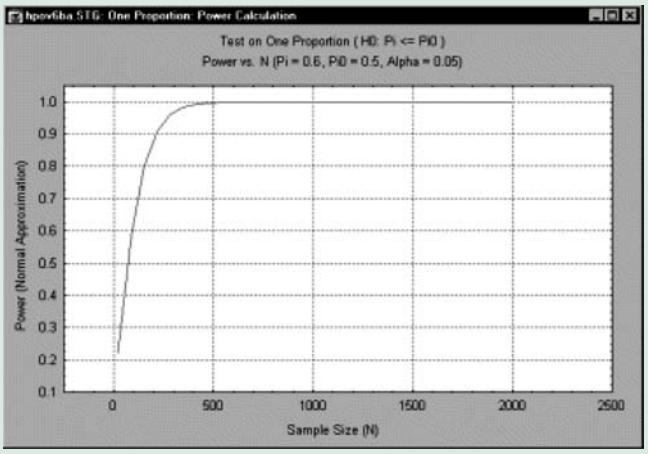
Researchers can make more precise estimates of the sample size required through a method called "statistical power analysis" (Kraemer & Thiemann 1987). Statistical power analysis requires a good advance estimate of the strength of the hypothesized relationship in the population. In addition, the math is complicated, so it helps to have some background in mathematics or to be able to consult a statistician. For these reasons, many researchers do not conduct formal power analyses when deciding how many cases to sample.

Exhibit 5.17 shows the results of a power analysis conducted to determine the sample size required to estimate a proportion in the population, when the null hypothesis is that that proportion is .50. For the sake of simplicity, it is assumed that the researcher wants to be 90% confident that the actual proportion differs from the null hypothesis of .50; in other words, he wants a sample size that will identify a difference from .5 that is significant at the .05 level. You can see that if the true proportion in the population is actually .55, a sample larger than 800 will be needed to detect this difference at the .05 level of significance. However, if the

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:30 PM Page 183

EXHIBIT 5.17 Power Graph



true proportion in the population is .60, then a random sample of only 200 cases is necessary. The required sample size falls off very gradually beyond this point, as the actual proportion in the population rises beyond .60.

It should be clear from Exhibit 5.17 that you must have a good estimate of the true population value of the statistic you are going to calculate. You also have to decide what significance level (such as .05) you want to achieve in your statistical test. Both of these factors can have a major impact on the number of cases you need to obtain.

You can obtain some general guidance about sample sizes from the current practices of social scientists. For professional studies of the national population in which only a simple description is desired, professional social science studies typically have used a sample size of between 1,000 and 1,500 people, with up to 2,500 being included if detailed analyses are planned. Studies of local or regional populations often sample only a few hundred people, in part because these studies lack sufficient funding to draw larger samples. Of course, the sampling error in these smaller studies is considerably larger than in a typical national study (Sudman 1976:87).

CONCLUSIONS

Sampling is a powerful tool for social science research. Probability sampling methods allow a researcher to use the laws of chance, or probability, to draw samples from which population parameters can be estimated with a high degree of confidence. A sample of just 1,000 or 1,500 individuals can be used to estimate reliably the characteristics of the population of a nation comprising millions of individuals.

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:30 PM Page 184

But researchers do not come by representative samples easily. Well-designed samples require careful planning, some advance knowledge about the population to be sampled, and adherence to systematic selection procedures—all so that the selection procedures are not biased. And even after the sample data are collected, the researcher's ability to generalize from the sample findings to the population is not completely certain. The best that he or she can do is to perform additional calculations that state the degree of confidence that can be placed in the sample statistic.

The alternatives to random, or probability-based, sampling methods are almost always much less palatable for quantitative studies, even though they are typically much cheaper. Without a method of selecting cases likely to represent the population in which the researcher is interested, research findings will have to be carefully qualified. Qualitative researchers whose goal is to understand a small group or setting in depth may necessarily have to use unrepresentative samples, but they must keep in mind that the generalizability of their findings will not be known. Additional procedures for sampling in qualitative studies will be introduced in Chapter 9.

Social scientists often seek to generalize their conclusions from the population that they studied to some larger target population. The validity of generalizations of this type is necessarily uncertain, because having a representative sample of a particular population does not at all ensure that what we find will hold true in other populations. Nonetheless, as you will see in Chapter 15, the cumulation of findings from studies based on local or otherwise unrepresentative populations can provide important information about broader populations.

KEY TERMS

Availability sampling
Census
Cluster
Cluster sampling
Disproportionate stratified sampling
Elements
Enumeration units
Inferential statistics
Nonprobability sampling method
Nonrespondents
Periodicity
Population
Population parameter
Probability of selection
Probability sampling method
Proportionate stratified sampling
Purposive sampling
Quota sampling
Random digit dialing
Random number table
Random sampling
Random sampling error
Replacement sampling
Representative sample
Sample
Sample statistic
Sampling error
Sampling frame
Sampling interval
Sampling units
Simple random sampling
Snowball sampling
Stratified random sampling
Systematic bias
Systematic random sampling
Target population

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:30 PM Page 185

HIGHLIGHTS

- Sampling theory focuses on the generalizability of descriptive findings to the population from which the sample was drawn. It also considers whether statements can be generalized from one population to another.
- Sampling is unnecessary when the elements that would be sampled are identical, but the complexity of the social world makes it difficult to argue very often that different elements are identical. Conducting a complete census of a population also eliminates the need for sampling, but the resources required for a complete census of a large population are usually prohibitive.
- Nonresponse undermines sample quality: It is the obtained sample, not the desired sample, that determines sample quality.
- Probability sampling methods rely on a random selection procedure to ensure no systematic bias in the selection of elements. In a probability sample, the odds of selecting elements are known, and the method of selection is carefully controlled.
- A sampling frame (a list of elements in the population) is required in most probability sampling methods. The adequacy of the sampling frame is an important determinant of sample quality.
- Simple random sampling and systematic random sampling are equivalent probability sampling methods in most of the situations. However, systematic random sampling is inappropriate for sampling from lists of elements that have a regular, periodic structure.
- Stratified random sampling uses prior information about a population to make sampling more efficient. Stratified sampling may be either proportionate or disproportionate. Disproportionate stratified sampling is useful when a research question focuses on a stratum or on strata that make up a small proportion of the population.
- Cluster sampling is less efficient than simple random sampling, but it is useful when a sampling frame is unavailable. It is also useful for large populations spread out across a wide area or among many organizations.
- Nonprobability sampling methods can be useful when random sampling is not possible, when a research question does not concern a larger population, and when a preliminary exploratory study is appropriate. However, the representativeness of nonprobability samples cannot be determined.
- The likely degree of error in an estimate of a population characteristic based on a probability sample decreases when the size of the sample and the homogeneity of the population from which the sample was selected increases. Sampling error is not affected by the proportion of the population that is sampled, except when that proportion is large. The degree of sampling error affecting a sample statistic can be estimated from the characteristics of the sample and knowledge of the properties of sampling distributions.

STUDENT STUDY SITE

To assist you in completing the Web exercises, please access the study site at www.pineforge.com/isw6 where you will find the Web exercises with accompanying links. You'll find other useful study materials such as self-quizzes and e-flashcards for each chapter, along with a group of carefully selected articles from research journals that illustrate the major concepts and techniques presented in the book.

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:31 PM Page 186

DISCUSSION QUESTIONS

1. When (if ever) is it reasonable to assume that a sample is not needed because "everyone is the same"—that is, the population is homogeneous? Does this apply to research such as that of Stanley Milgram's on obedience to authority? What about investigations of student substance abuse? How about investigations of how people (or their bodies) react to alcohol? What about research on likelihood of voting (the focus of Chapter 14)?
2. All adult U.S. citizens are required to participate in the decennial census, but some do not. Some social scientists have argued for putting more resources into a large representative sample, so that more resources are available to secure higher rates of response from hard-to-include groups. Do you think that the U.S. Census should shift to a probability-based sampling design? Why or why not?
3. What increases sampling error in probability-based sampling designs? Stratified rather than simple random sampling? Disproportionate (rather than proportionate) stratified random sampling? Stratified rather than cluster random sampling? Why do researchers select *disproportionate* (rather than proportionate) stratified samples? Why do they select cluster rather than simple random samples?
4. What are the advantages and disadvantages of probability-based sampling designs compared with nonprobability-based designs? Could any of the researches that are described in this chapter with a nonprobability-based design have been conducted instead with a probability-based design? What are the difficulties that might have been encountered in an attempt to use random selection? How would you discuss the degree of confidence you can place in the results obtained from research using a nonprobability-based sampling design?

PRACTICE EXERCISES

1. Select a random sample using the table of random numbers in Appendix E. Compute a statistic based on your sample, and compare it with the corresponding figure for the entire population. Here's how to proceed:
   a. First, select a very small population for which you have a reasonably complete sampling frame. One possibility would be the list of asking prices for houses advertised in your local paper. Another would be the listing of some characteristic of states in a U.S. Census Bureau publication, such as average income or population size.
   b. The next step is to create your sampling frame, a numbered list of all the elements in the population. If you are using a complete listing of all elements, as from a U.S. Census Bureau publication, the sampling frame is the same as the list. Just number the elements (states). If your population is composed of housing ads in the local paper, your sampling frame will be those ads that contain a housing price. Identify these ads, and then number them sequentially, starting with 1.
   c. Decide on a method of picking numbers out of the random number table in Appendix E, such as taking every number in each row, row by row (or you may move down or diagonally across the columns). Use only the first (or last) digit in each number if you need to select 1 to 9 cases or only the first (or last) two digits if you want fewer than 100 cases.
   d. Pick a starting location in the random number table. It's important to pick a starting point in an unbiased way, perhaps by closing your eyes and then pointing to some part of the page.

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:31 PM Page 187

e. Record the numbers you encounter as you move from the starting location in the direction you decided on in advance, until you have recorded as many random numbers as the number of cases you need in the sample. If you are selecting states, 10 might be a good number. Ignore numbers that are too large (or too small) for the range of numbers used to identify the elements in the population. Discard duplicate numbers.

f. Calculate the average value in your sample for some variable that was measured—for example, population size in a sample of states or housing price for the housing ads. Calculate the average by adding up the values of all the elements in the sample and dividing by the number of elements in the sample.

g. Go back to the sampling frame and calculate this same average for all elements in the list. How close is the sample average to the population average?

h. Estimate the range of sample averages that would be likely to include 90% of the possible samples.

2. Draw a snowball sample of people who are involved in bungee jumping or some other uncommon sport that does not involve teams. Ask friends and relatives to locate a first contact, and then call or visit this person and ask for names of others. Stop when you have identified a sample of 10. Review the problems you encountered, and consider how you would proceed if you had to draw a larger sample.

3. Two lesson sets on the study site will help you review the terminology involved in "Identifying Sampling Techniques" and the logic of "Assessing Generalizability."

4. Identify one article at the book's study site, www.pineforge.com/isw6/learning.htm that used a survey research design. Describe the sampling procedure. What type was it? Why did the author(s) use this particular type of sample?

ETHICS QUESTIONS

1. How much pressure is too much pressure to participate in a probability-based sample survey? Is it OK for the U.S. government to mandate legally that all citizens participate in the decennial census? Should companies be able to require employees to participate survey research about work-related issues? Should students be required to participate in surveys about teacher performance? Should parents be required to consent to the participation of their high school–age students in a survey about substance abuse and health issues? Is it OK to give monetary incentives for participation in a survey of homeless shelter clients? Can monetary incentives be coercive? Explain your decisions.

2. Federal regulations require special safeguards for research on persons with impaired cognitive capacity. Special safeguards are also required for research on prisoners and on children. Do you think special safeguards are necessary? Why or why not? Do you think it is possible for individuals in any of these groups to give "voluntary consent" to research participation? What procedures might help make consent to research truly voluntary in these situations? How could these procedures influence sampling plans and results?

WEB EXERCISES

1. Research on homelessness has been rising in recent years as housing affordability has declined. Search the Web for sites that include the word *homelessness* and see what you find. You might try limiting your search to those that also contain the word *census*. Pick a site and write a paragraph about what you learned from it.

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:31 PM Page 188

2. Check out the "people and households" section of the U.S. Bureau of the Census Web site: www.census.gov. Based on some of the data you find there, write a brief summary of some aspects of the current characteristics of the American population.

SPSS EXERCISES

1. Take a look again at the distribution of support for capital punishment (CAPPUN), this time with what is called a "frequency distribution."
   a. Click Analyze\Descriptive Statistics\Frequencies.
   b. Highlight CAPPUN and click on the arrow that sends it over to the Variables window, then click OK.

   Examine the percentages in the Valid percent column. What percentage of the American population in 2006 favored capital punishment?

2. Now select random samples of the GSS2006x respondents and see how the distribution of CAPPUN in these subsamples compares with that for the total GSS sample:
   a. Go to the Data Editor window, and select a random sample containing 40 of the respondents. From the menu:

1. Click Data\Select cases\All Cases\OK.
2. Click Select cases\Random sample of cases\Sample.
3. Select exactly 40 cases from the first 100 cases.
4. Click Continue\OK. (Before you click OK, be sure that the "Filter out unselected cases" box is checked.)
   a. Determine the percentage of the subsample that favored capital punishment by repeating the steps in SPSS Exercise 1. Record the subsample characteristics and its percentage.
   b. Now, repeat Steps 2a and 2b 10 times. Each time, add 100 to the "first 100 cases" request (so that on the last step you will be requesting "Exactly 40 cases from the first 1,000 cases").
   c. Select a random sample containing five of the respondents. Now repeat Steps 2a through 2c, this time for samples of five.
   d. Plot the results of Steps 2c and 2d on separate sheets of graph paper. Each graph's horizontal axis will represent the possible range of percentages (from 0 to 100, perhaps in increments of 5); the vertical axis will represent the number of samples in each range of percentages (perhaps ranging from 0 to 10). Make an X to indicate this percentage for each sample. If two samples have the same percentage, place the corresponding Xs on top of each other. The X for each sample should be one unit high on the vertical axis.
   e. Draw a vertical line corresponding to the point on the horizontal axis that indicates the percentage of the total GSS sample that favors capital punishment.
   f. Describe the shape of both the graphs. These are the sampling distributions for the two sets of samples. Compare them with each other. Do the percentages from the larger samples tend to be closer to the mean of the entire sample (as obtained in SPSS Exercise 1)? What does this tell you about the relationship between sample size and sampling error?

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE

05-Schutt 6e-45771:FM-Schutt5e(4853) (for student CD).qxd 9/29/2008 11:31 PM Page 189

DEVELOPING A RESEARCH PROPOSAL

Consider the possibilities for sampling (Exhibit 2.12, #8).

1. Propose a sampling design that would be appropriate if you were to survey students on your campus only. Define the population, identify the sampling frame(s), and specify the elements and any other units at different stages. Indicate the exact procedure for selecting people to be included in the sample.
2. Propose a different sampling design for conducting your survey in a larger population, such as your city, state, or the entire nation.

Unproofed pages. Not to be sold, copied, or redistributed. Property of SAGE