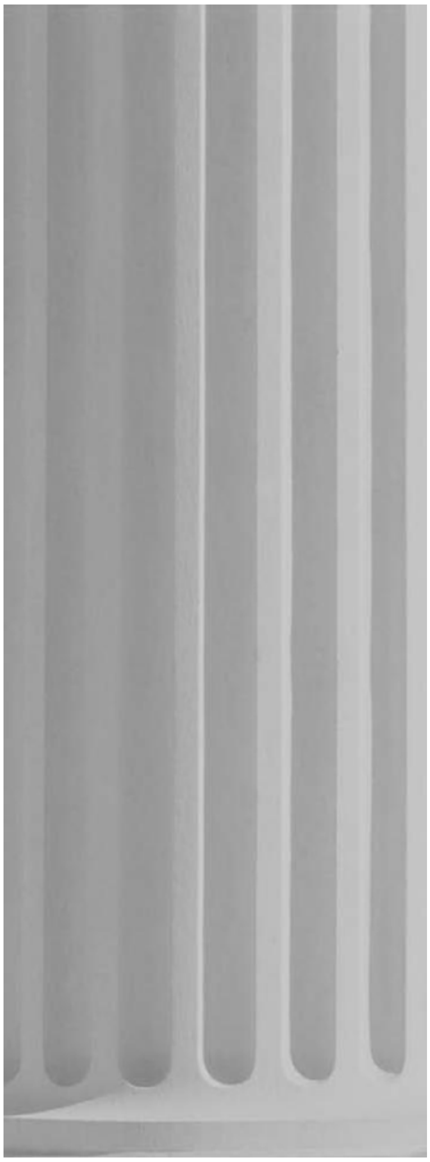
CHAPTER 2

Sampling



KEY TERMS

bell curve
cluster or area random sampling
expert sampling
external validity
gradient of similarity
heterogeneity sampling
modal instance sampling
multistage sampling
nonprobability sampling
nonproportional quota sampling
population
population parameter
probability sampling
proportional quota sampling
proximal similarity model
quota sampling
random selection
response
sampling
sampling distribution
sampling error
sampling frame
sampling model
simple random sampling
snowball sampling
standard deviation
standard error
statistic
stratified random sampling
systematic random sampling
validity

OUTLINE

Sampling is the process of selecting units (such as people and organizations) from a **population** of interest so that by studying the **sample** you can fairly generalize your results to the population from which the units were chosen. In this chapter, I begin by covering some of the key terms in sampling like population and **sampling frame**. Then, because some types of sampling rely on quantitative models, I'll talk about some of the statistical terms used in sampling. Finally, I'll discuss the major distinction between probability and nonprobability sampling methods and work through the major types in each.

population
The group you want to generalize to and the group you sample from in a study.

sample
The actual units you select to participate in your study.

sampling frame
The list from which you draw your sample. In some cases, there is no list; you draw your sample based upon an explicit rule. For instance, when doing quota sampling of passersby at the local mall, you do not have a list per se, and the sampling frame consists of both the population of people who pass by within the time frame

2-1 External Validity

It may seem odd to begin a chapter about sampling by discussing **external validity**. So, I want to take a minute to explain how these topics are linked. In research, our sample consists of the people who actually participate in our study. But when we conduct research, we are often interested not just in reaching conclusions about our sample in the time and place where we conducted our study but also in making some conclusions that are broader than that, in concluding what might happen with other people at other times and in other places than just our sample. When we try to reach conclusions that extend beyond the sample in our study, we say that we are generalizing. So, why is external validity so important to the issue of sampling? Because external validity is centrally related to the idea of generalizing. Recall from Section 1-3, Validity of Research, in Chapter 1, that **validity** of any type refers to the approximate truth of propositions, inferences, or conclusions. *External validity* refers to the approximate truth of conclusions that involve generalizations, or more broadly, the **generalizability** of conclusions. Put in everyday terms, *external validity* is the degree to which the conclusions in your study would hold for other persons in other places and at other times (Cook & Campbell, 1979; Shadish, Cook, & Campbell, 2002).

external validity
The degree to which the conclusions in your study would hold for other persons in other places and at other times.

validity
The best available approximation of the truth of a given proposition, inference, or conclusion.

generalizability
The degree to which study conclusions are valid for members of the population not included in the study sample.

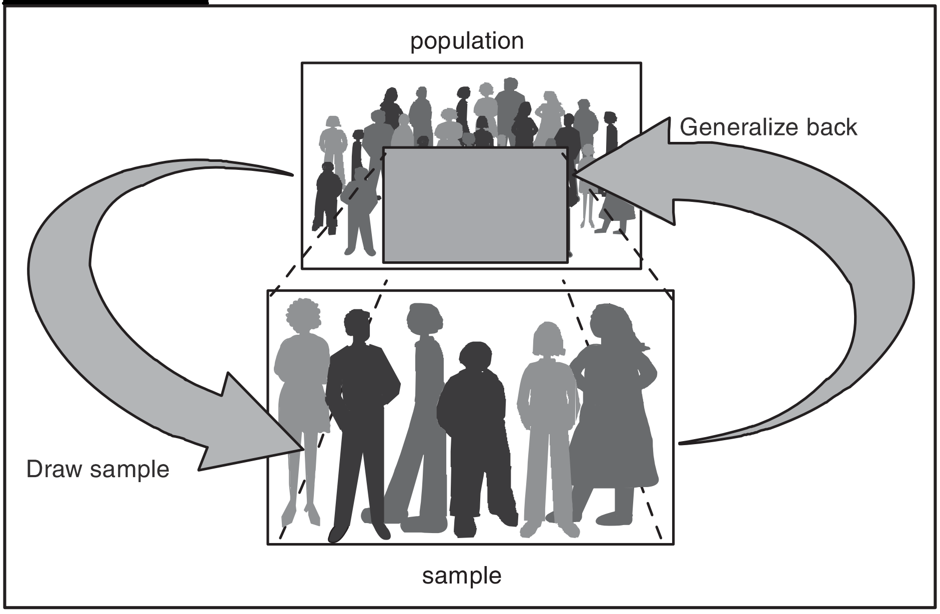
Science approaches the providing of evidence for a generalization in two major ways. I'll call the first approach the **sampling model**. In the sampling model, you start by identifying the population you would like to generalize to (Figure 2–1). Then, you draw a fair sample from that population and conduct your research with the sample. Finally, because the sample is representative of the population, you can automatically generalize your results back to the population. This approach has several problems. First, at the time of your study, you might not know what part of the population you will ultimately want to generalize to. Second, you may not be able to draw a fair or representative sample easily. Third, it's impossible to sample across all times that you might like to generalize to, such as next year.

sampling model
A model for generalizing in which you identify your population, draw a fair sample, conduct your research, and finally, generalize your results to other population groups.

The second approach to generalizing is called the **proximal similarity model** (Figure 2–2). *Proximal* means nearby and *similarity* means ... well, it means

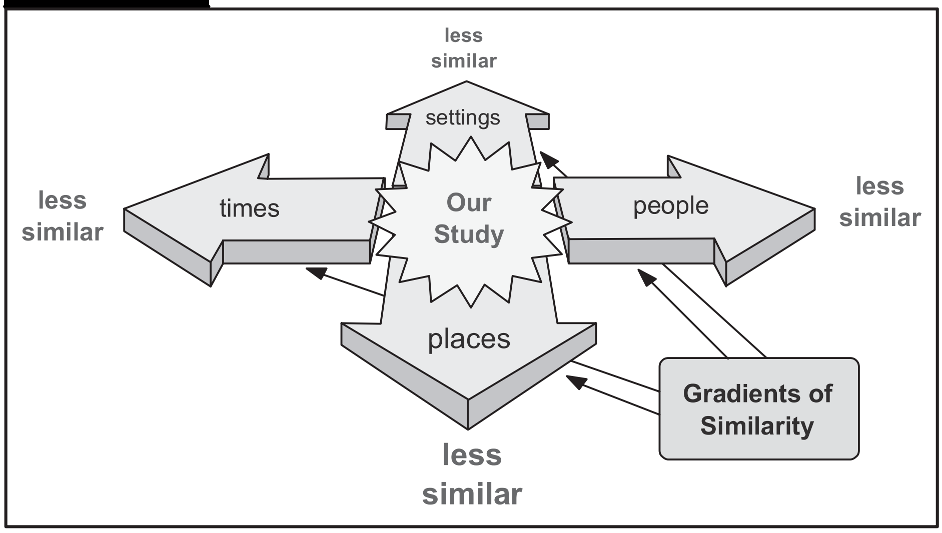
proximal similarity model
A model for generalizing from your study to another context based upon the degree to which the other context is similar to your study context.

FIGURE 2–1 **The sampling model for external validity**



The researcher draws a sample for a study from a defined population to generalize the results to the population.

FIGURE 2–2 **The proximal similarity model for external validity**



similarity. The term *proximal similarity* was suggested by Donald T. Campbell as an appropriate relabeling of the term *external validity* (although he was the first to admit that it probably wouldn't catch on). With proximal similarity, you begin by thinking about different generalizability contexts and developing a theory about which contexts are more like your study and which are less so. For instance, you might imagine several settings that have people who are more similar to the people in your study or people who are less similar. This process also holds for times and places. When you place different contexts in terms of their relative similarities, you can call this implicit theoretical dimension a **gradient of similarity**. After you develop this proximal similarity framework, you can generalize. How? You can generalize the results of your study to other persons, places, or times that are more like (that is, more proximally similar to) your study. Notice that here, you can never generalize with certainty; these generalizations are always a question of more or less similar.

gradient of similarity
The dimension along which your study context can be related to other potential contexts to which you might wish to generalize. Contexts that are closer to yours along the gradient of similarity of place, time, people, and so on can be generalized to with more confidence than ones that are further away.

2-1a Threats to External Validity

A threat to external validity is an explanation of how you might be wrong in making a generalization. For instance, imagine that you conclude that the results of your study (which was done in a specific place, with certain types of people, and at a specific time) can be generalized to another context (for instance, another place, with slightly different people, at a slightly later time). In such a case, three major threats to external validity exist because there are three ways you could be wrong: people, places, and times. Your critics could, for example, argue that the results of your study were due to the unusual type of people who were in the study, or, they could claim that your results were obtained only because of the unusual place in which you performed the study. (Perhaps you did your educational study in a college town with lots of high-achieving, educationally oriented kids.) They might suggest that you did your study at a peculiar time. For instance, if you did your smoking-cessation study the week after the Surgeon General issued the well-publicized results of the latest smoking and cancer studies, you might get different results than if you had done it the week before.

2-1b Improving External Validity

random selection
Process or procedure that assures that the different units in your population are selected by chance.

How can you improve external validity? One way, based on the sampling model, suggests that you do a good job of drawing a sample from a population. For instance, you should use **random selection**, if possible, rather than a nonrandom procedure. In addition, once selected, you should try to ensure that the respondents participate in your study and that you keep your dropout rates low. A second approach would be to use the theory of proximal similarity more effectively. How? Perhaps you could do a better job of describing the ways your contexts differ from others by providing data about the degree of similarity between various groups of people, places, and even times. You might even be able to map out the degree of proximal similarity among various contexts with a methodology like concept mapping as discussed in Chapter 1. Perhaps the best approach to criticisms of generalizations is simply to show critics that they're wrong—do your study in a variety of places, with different people, and at different times. That is, your external validity (ability to generalize) will be stronger the more you replicate your study.

2-2 Sampling Terminology

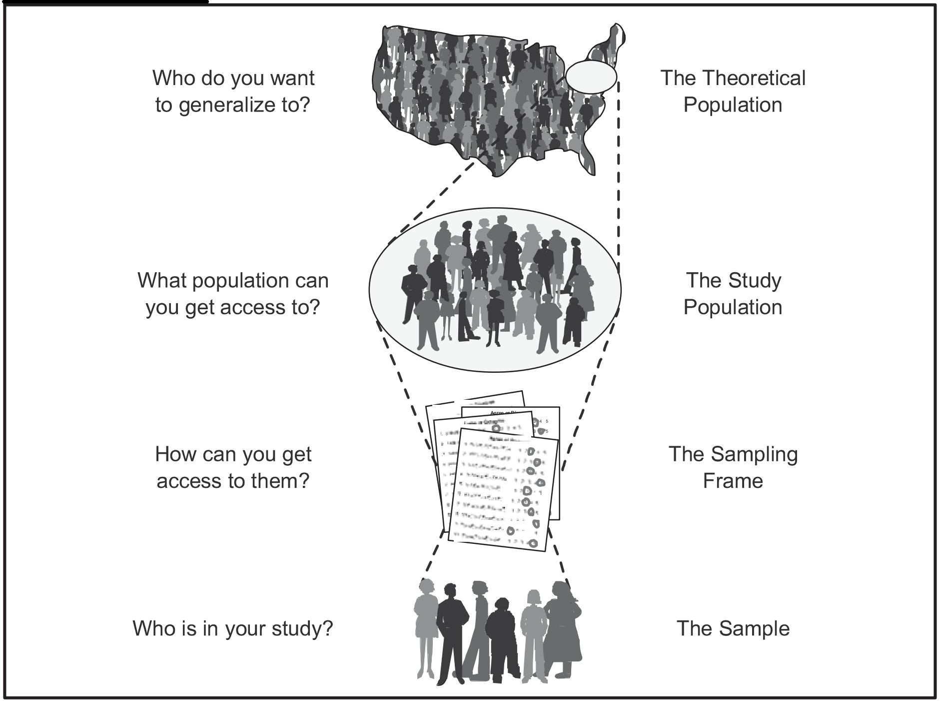
As with anything else in life you have to learn the language of an area if you're going to ever hope to use it. Here, I want to introduce several different terms for the major groups that are involved in a sampling process and the role that each group plays in the logic of sampling.

The major question that motivates sampling in the first place is: "Whom do you want to generalize to?" (Or should it be: "To whom do you want to generalize?") In most social research, you are interested in more than just the people directly participating in your study. You would like to be able to talk in general terms and not be confined to only the people in your study. Now, at times you won't be concerned about generalizing. Maybe you're just evaluating a program in a local agency and don't care whether the program would work with other people in other places and at other times. In that case, sampling and generalizing might not be of interest. In other cases, you would really like to be able to generalize almost universally.

census
A kind of survey that involves a complete enumeration of the entire population of interest.

Indeed, sometimes an actual **census** is needed. Taking a census means a counting or enumeration of every member of the population. For example, one of the main tasks undertaken by the U.S. Census Bureau is to periodically compile a record of every person living in the United States. This complete listing is necessary for the proper conduct of elections (for example, the electoral college used in

FIGURE 2–3 **The different groups in the sampling model**



presidential elections) as well as taxation and other functions of government in which a complete record rather than an estimate is needed. (By the way the U.S. Census Bureau is a gold mine of interesting information on many aspects of the U.S. population: http://www.census.gov.)

When psychologists do research, they are often interested in developing theories that would hold for all humans; in contrast, in most applied social research, researchers are interested in generalizing to specific groups. The group you wish to generalize to is called the *population* in your study (Figure 2–3). This is the group you would like to sample from because this is the group you are interested in generalizing to. Let's imagine that you want to generalize to urban homeless males between the ages of 30 and 50 in the United States. If that is the population of interest, you are likely to have a hard time developing a reasonable sampling plan. You are probably not going to find an accurate listing of this population, and even if you did, you would almost certainly not be able to mount a national sample across hundreds of urban areas. So you probably should make a distinction between the population you would like to generalize to, and the population that is accessible to you. We'll call the former the *theoretical population* and the latter the *accessible population*. In this example, the accessible population might be homeless males between the ages of 30 and 50 in six selected urban areas across the United States.

After you identify the theoretical and accessible populations, you have to do one more thing before you can actually draw a sample: get a list of the members of the accessible population. (Or, you have to spell out in detail how you will contact them to ensure representativeness.) The listing of the accessible population from which you'll draw your sample is called the *sampling frame*. If you were doing a phone survey and selecting names from the telephone book, the phone book would be your sampling frame. That wouldn't be a great way to sample because significant subportions of the population either don't have a phone or have moved in or out of the area since the last phone book was printed. Notice that in this case, you might identify the area code and all three-digit prefixes within that area code and draw a sample simply by randomly dialing numbers (cleverly known as *random-digit-dialing*). In this case, the sampling frame is not a list *per se*, but is rather a procedure that you follow as the actual basis for sampling. Finally, you actually draw your sample (using one of the many sampling procedures described later in this chapter). The *sample* is the group of people you select to be in your study. Notice that I didn't say that the sample was the group of people who are actually in your study. You may not be able to contact or recruit all of the people you actually sample, or some could drop out over the course of the study. The group that actually completes your study is a subsample of the sample; it doesn't include nonrespondents or dropouts. (The problem of nonresponse and its effects on a study will be addressed in Chapter 7 when discussing mortality threats to internal validity.)

People often confuse the idea of random selection with the idea of random assignment. You should make sure that you understand the distinction between random selection and random assignment described in Chapter 8.

At this point, you should appreciate that sampling is a difficult multistep process and that you can go wrong in many places. In fact, as you move from each step to the next in identifying a sample, there is the possibility of introducing systematic error or *bias*. For instance, even if you are able to identify perfectly the population of interest, you may not have access to all of it. Even if you do, you may not have a complete and accurate enumeration or sampling frame from which to select. Even if you do, you may not draw the sample correctly or accurately. And, even if you do, your participants may not all come and they may not all stay. Depressed yet? Sampling is a difficult business indeed. At times like this, I'm reminded of what one of my professors, Donald Campbell, used to say (I'll paraphrase here): "Cousins to the amoeba, it's amazing that we know anything at all!"

2-3 Statistical Terms in Sampling

response
A specific measurement value that a sampling unit supplies.

statistic
A specific value that is estimated from data.

population parameter
The mean or average you would obtain if you were able to sample the entire population.

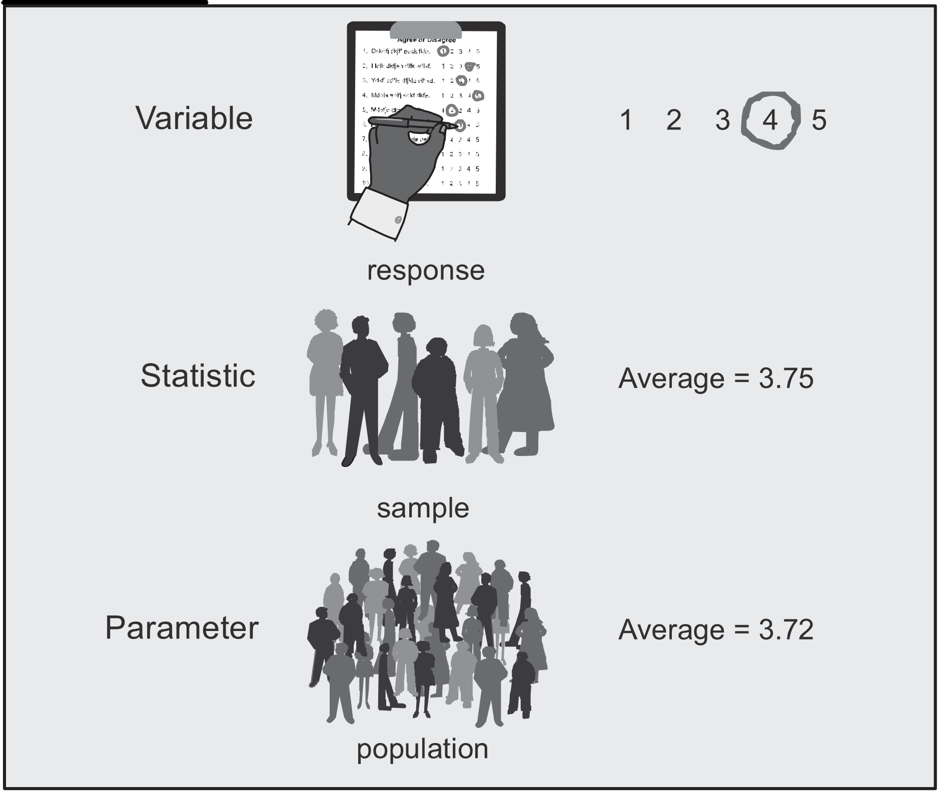
Let's begin by defining some simple terms that are relevant here. First, let's look at the results of sampling efforts. When you sample, the units that you sample—usually people—supply you with one or more responses. In this sense, a **response** is a specific measurement value that a sampling unit supplies. In Figure 2–4, the person responding to a survey instrument gives a response of "4." When you look across the responses for your entire sample, you use a **statistic**. You can use a variety of statistics: mean, median, mode, and so on. In the Analysis chapter of this book (Chapter 12), you'll find more detailed information on the most commonly used statistics along with some more elaborate examples. In this example, the mean or average for the sample is 3.72; but the reason you sample is to get an estimate for the population from which you sampled. If you could, you would probably prefer to measure the entire population. If you measure the entire population and calculate a value like a mean or average, this is not referred to as a *statistic*; it is a **population parameter**.

2-3a The Sampling Distribution

sampling distribution
The theoretical distribution of an infinite number of samples of the population of interest in your study.

So how do you get from sample statistic to an estimate of the population parameter? A crucial midway concept you need to understand is the **sampling distribution**. To understand it, you have to be able and willing to do a thought experiment. Imagine that instead of just taking a single sample like you do in a typical study, you took three independent samples of the same population. Furthermore, imagine that for each of your three samples, you collected a single response and computed a single statistic, say, the mean of the response for each sample. This is depicted in the top part of Figure 2–5. Even though all three samples came from the same population, you wouldn't expect to get the exact same statistic from each. They would differ slightly due to the random luck of the draw or to the natural fluctuations or vagaries of drawing a sample. However, you would expect all three samples to yield a similar statistical estimate because they were drawn from the same population.

FIGURE 2–4 **Statistical terms in sampling**



Now, for the leap of imagination! Imagine that you took an *infinite* number of samples from the same population and computed the average for each one. If you plotted the averages on a histogram or bar graph, you should find that most of them converge on the same central value and that you get fewer and fewer samples that have averages farther up or down from that central value. In other words, the bar graph would be well described by the **bell curve** shape that is an indication of a normal distribution in statistics. This is depicted in the bottom part of Figure 2–5. The distribution of an infinite number of samples of the same size as the sample in your study is known as the sampling distribution.

bell curve
Smoothed histogram or bar graph describing the expected frequency for each value of a variable. The name comes from the fact that such a distribution often has the shape of a bell.

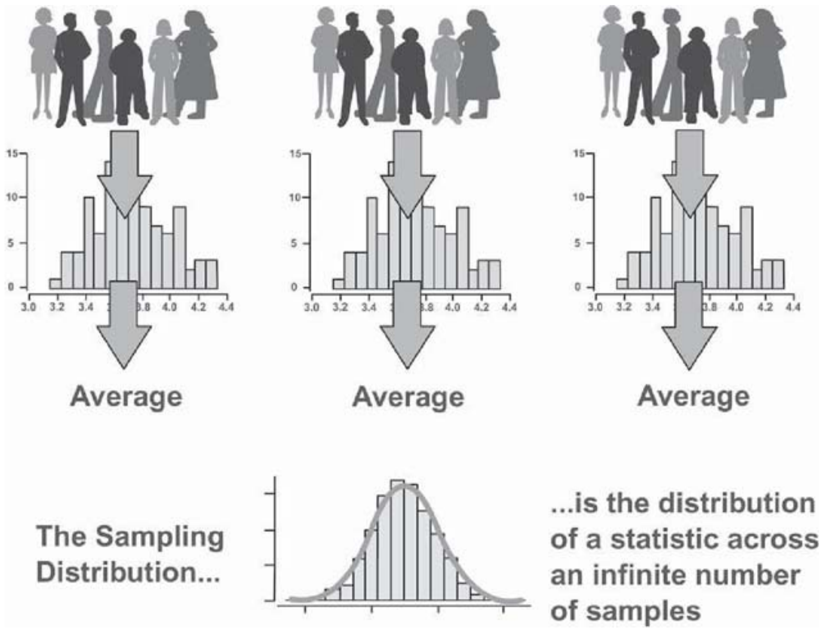
You don't ever actually construct a sampling distribution. Why not? You're not paying attention! Because to construct one, you would have to take an *infinite* number of samples and at least the last time I checked, on this planet infinite is not a number we know how to reach. So why do researchers even talk about a sampling distribution? Now that's a good question! Because you need to realize that your sample is just one of a potentially infinite number of samples that you could have taken. When you keep the sampling distribution in mind, you realize that while the statistic from your sample is probably near the center of the sampling distribution (because most of the samples would be there) you could have gotten one of the extreme samples just through the luck of the draw. If you take the average of the sampling distribution—the average of the averages of an infinite number of samples—you would be much closer to the true population average—the parameter of interest.

So the average of the sampling distribution is essentially equivalent to the parameter. But what is the **standard deviation** of the sampling distribution? (Okay, don't remember what a standard deviation is? This is discussed in detail in the section "Descriptive Statistics" in Chapter 10, "Analysis.") The standard deviation of the sampling distribution tells us something about how different samples would be distributed. In statistics it is referred to as the **standard error** (so you can keep it

standard deviation
The spread or variability of the scores around their average in a *single sample.* The standard deviation, often abbreviated SD, is mathematically the square root of the variance. The standard deviation and variance both measure dispersion, but because the standard deviation is measured in the same units as the original measure and the variance is measured in squared units, the standard deviation is usually more directly interpretable and meaningful.

standard error
The spread of the averages around the average of averages in a sampling distribution.

FIGURE 2–5 **The sampling distribution**



separate in your minds from standard deviations. Getting confused? Go get a cup of coffee and come back in 10 minutes . . .. Okay, let's try once more . . .). A *standard deviation* is the spread of the scores around the average in a *single sample*. The standard error is the spread of the averages around the average of averages in a *sampling distribution*. Got it?

2-3b Sampling Error

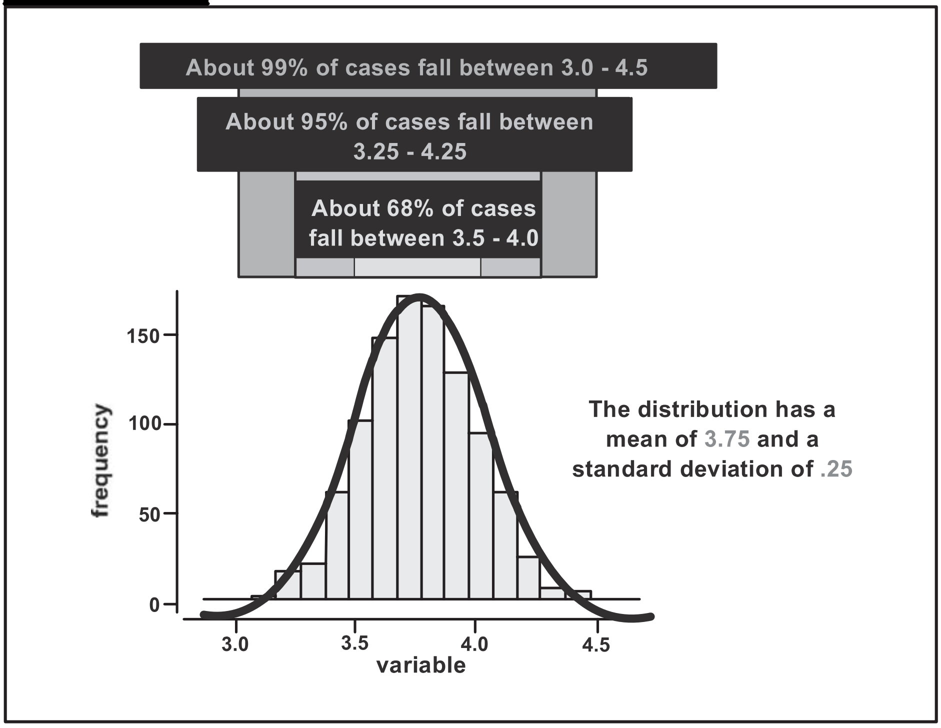
sampling error
The error in measurement associated with sampling.

In sampling, the standard error is called **sampling error**. Sampling error gives you some idea of the precision of your statistical estimate. A low sampling error means that you had relatively less variability or range in the sampling distribution. But here I go again; you never actually see the sampling distribution! So how do you calculate sampling error? You base your calculation *on the standard deviation of your sample*: the greater the sample's standard deviation, the greater the standard error (and the sampling error). The standard error is also related to the sample size: the greater your sample size, the *smaller* the standard error. Why? Because the greater the sample size, the closer your sample is to the actual population itself. If you take a sample that consists of the entire population, you actually have no sampling error because you don't have a sample; you have the entire population (that is, a census). In that case, the mean you estimate is the parameter.

2-3c The 68, 95, 99 Percent Rule

You've probably heard this one before, but it's so important that it's always worth repeating. There is a general rule that applies whenever you have a normal or bell-shaped distribution. Start with the average—the center of the distribution. If you go up and down (that is, left and right) one standard unit, you will include approximately 68 percent of the cases in the distribution (68 percent of the area under the curve). If you go up and down two standard units, you will include approximately 95 percent of the cases. If you go plus or minus three standard units, you will include 99 percent of the cases.

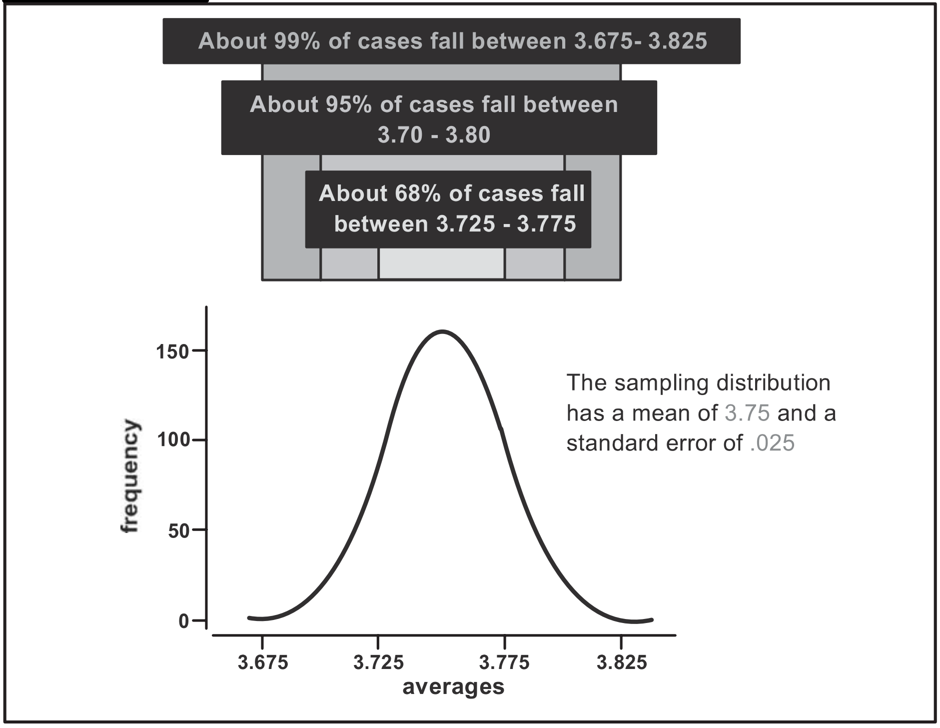
FIGURE 2–6 **The 68, 95, 99 percent rule**



Notice that I didn't specify in the previous few sentences whether I was talking about standard *deviation* units or standard *error* units. That's because the same rule holds for both types of distributions (the raw data and sampling distributions). For instance, in Figure 2–6, the mean of the distribution is 3.75 and the standard unit is .25. (If this were a distribution of raw data, we would be talking in standard-deviation units. If it were a sampling distribution, we'd be talking in standard-error units.) If you go up and down one standard unit from the mean, you would be going up and down .25 from the mean of 3.75. Within this range—3.5 to 4.0—you would expect to see approximately 68 percent of the cases. This section is marked in red on Figure 2–6. I leave it to you to figure out the other ranges. What does this all mean, you ask. If you are dealing with raw data and you know the mean and standard deviation of a sample, you can *predict* the intervals within which 68, 95, and 99 percent of your cases would be expected to fall. We call these intervals the—guess what—68, 95, and 99 percent confidence intervals.

Now, here's where everything should come together in one great aha! experience if you've been following along. If you have a *sampling distribution*, you should be able to predict the 68, 95, and 99 percent confidence intervals for where the population parameter should be; and isn't that why you sampled in the first place? So that you could predict where the population is on that variable? There's only one hitch. You don't actually have the sampling distribution. (I know this is the third time I've said this.) However, you do have the distribution for the sample itself, and from that distribution, you can estimate the standard error (the sampling error) because it is based on the standard deviation and you have that. Of course, you don't actually know the population parameter value; you're trying to find that out, but you can use your best estimate for that—the sample statistic. Now, if you have the mean of the sampling distribution (or set it to the mean from your sample) and you have an estimate of the standard error, which you calculate that from your sample, you have the two key ingredients that you need for your sampling distribution to estimate confidence intervals for the population parameter.

FIGURE 2–7 **Estimating the population using a sampling distribution**



Perhaps an example will help. Let's assume you did a study and drew a single sample from the population. Furthermore, let's assume that the average for the sample was 3.75 and the standard deviation was .25. This is the raw data distribution depicted in Figure 2–7. What would the sampling distribution be in this case? Well, you don't actually construct it (because you would need to take an infinite number of samples) but you *can* estimate it. For starters, you must assume that the mean of the sampling distribution is the mean of the sample, which is 3.75. Then, you calculate the standard error. To do this, use the standard deviation for your sample and the sample size (in this case $N = 100$), which gives you a standard error of .025 (just trust me on this). Now you have everything you need to estimate a confidence interval for the population parameter. You would estimate that the probability is 68 percent that the true parameter value falls between 3.725 and 3.775 (3.75 plus and minus .025); that the 95 percent confidence interval is 3.700 to 3.800; and that you can say with 99 percent confidence that the population value is between 3.675 and 3.825. Using your sample, you have just estimated the average for your population (that is, the mean of the sample which is 3.75) and you have given odds that the actual population mean falls within certain ranges.

2-4 Probability Sampling

probability sampling
Method of sampling that utilizes some form of random selection.

A **probability sampling** method is any method of sampling that utilizes some form of random selection. To have a random selection method, you must set up some process or procedure that ensures that the different units in your population have equal probabilities of being chosen. Humans have long practiced various forms of random selection, such as picking a name out of a hat, or choosing the short straw. These days, we tend to use computers as the mechanism for generating random numbers as the basis for random selection.

2-4a Some Definitions

Before I can explain the various probability methods, I have to define the following basic terms:

- N is the number of cases in the sampling frame.
- n is the number of cases in the sample.
- ${}_NC_n$ is the number of combinations (subsets) of n from N.
- $f = n/N$ is the sampling fraction.

That's it. Now that you understand those terms, I can define the different probability sampling methods.

2-4b Simple Random Sampling

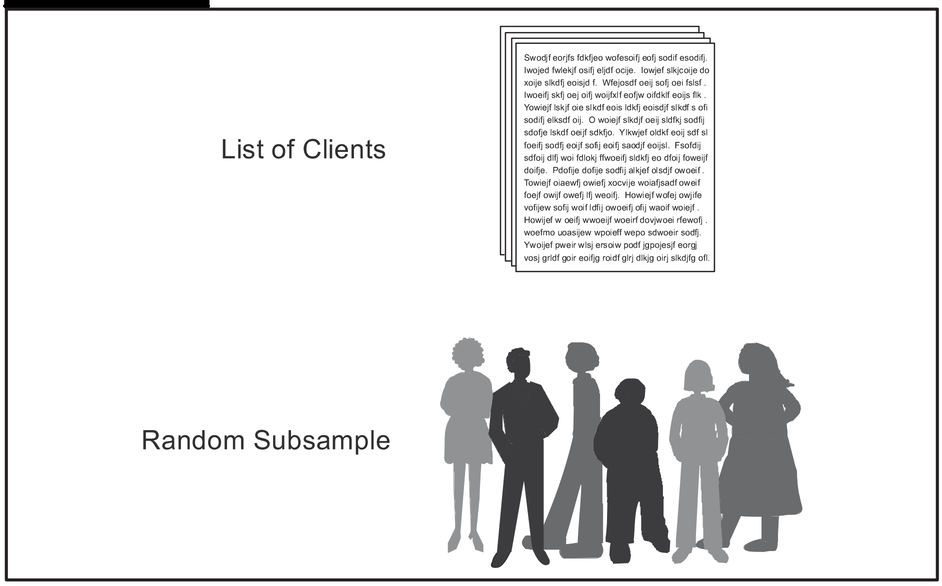
The simplest form of random sampling is called **simple random sampling**. Pretty tricky, huh? Here's the quick description of simple random sampling:

- **Objective:** To select n units out of N such that each ${}_NC_n$ has an equal chance of being selected.
- **Procedure:** Use a table of random numbers, a computer random-number generator, or a mechanical device to select the sample.

simple random sampling
A method of sampling that involves drawing a sample from a population so that every possible sample has an equal probability of being selected.

Let's see if I can make this somewhat stilted description a little more real. How do you select a simple random sample? Let's assume that you are doing some research with a small service agency to assess clients' views of quality of service over the past year. First, you have to get the sampling frame organized. To accomplish this, you go through agency records to identify every client over the past 12 months. If you're lucky, the agency has accurate computerized records and can quickly produce such a list (Figure 2–8). Then, you have to draw the *sample* and decide on the number of clients you would like to have in the final sample. For the sake of the example, let's say you want to select 100 clients to survey and that there were 1000 clients over the past 12 months. Then, the sampling fraction is $f = n/N = 100/1000 = .10$, or 10 percent. To draw the sample, you have several options. You could print the list of 1000 clients, tear them into separate strips, put the strips in a hat, mix them up, close your eyes, and pull out the first 100. This mechanical procedure would be tedious and the quality of the sample would depend on how thoroughly you mixed up the paper

FIGURE 2–8 **Simple random sampling**



strips and how randomly you reached into the hat. Perhaps a better procedure would be to use the kind of ball machine that is popular with many of the state lotteries. You would need three sets of balls numbered 0 to 9, one set for each of the digits from 000 to 999. (If you select 000 you call that 1000.) Number the list of names from 1 to 1000 and then use the ball machine to select the three digits that selects each person. The obvious disadvantage here is that you need to get the ball machines. (Where do they make those things, anyway? Is there a ball machine industry?)

Neither of these mechanical procedures is feasible and, with the development of inexpensive computers, there is a much easier way. Here's a simple procedure that's especially useful if you have the names of the clients already on the computer. Many computer programs can generate a series of random numbers. Let's assume you copy and paste the list of client names into a column in an Excel spreadsheet. Then, in the column right next to it paste the function = RAND(), which is Excel's way of putting a random number between 0 and 1 in the cells. Then, sort both columns—the list of names and the random number—by the random numbers. This rearranges the list in random order from the lowest to the highest random number. Then, all you have to do is take the first hundred names in this sorted list. Pretty simple. You could probably accomplish the whole thing in under a minute.

Simple random sampling is easy to accomplish and explain to others. Because simple random sampling is a fair way to select a sample, it is reasonable to generalize the results from the sample back to the population. Simple random sampling is not the most statistically efficient method of sampling and you may—just because of the luck of the draw—not get a good representation of subgroups in a population. To deal with these issues, you have to turn to other sampling methods.

2-4c Stratified Random Sampling

stratified random sampling
A method of sampling that involves dividing your population into homogeneous subgroups and then taking a simple random sample in each subgroup.

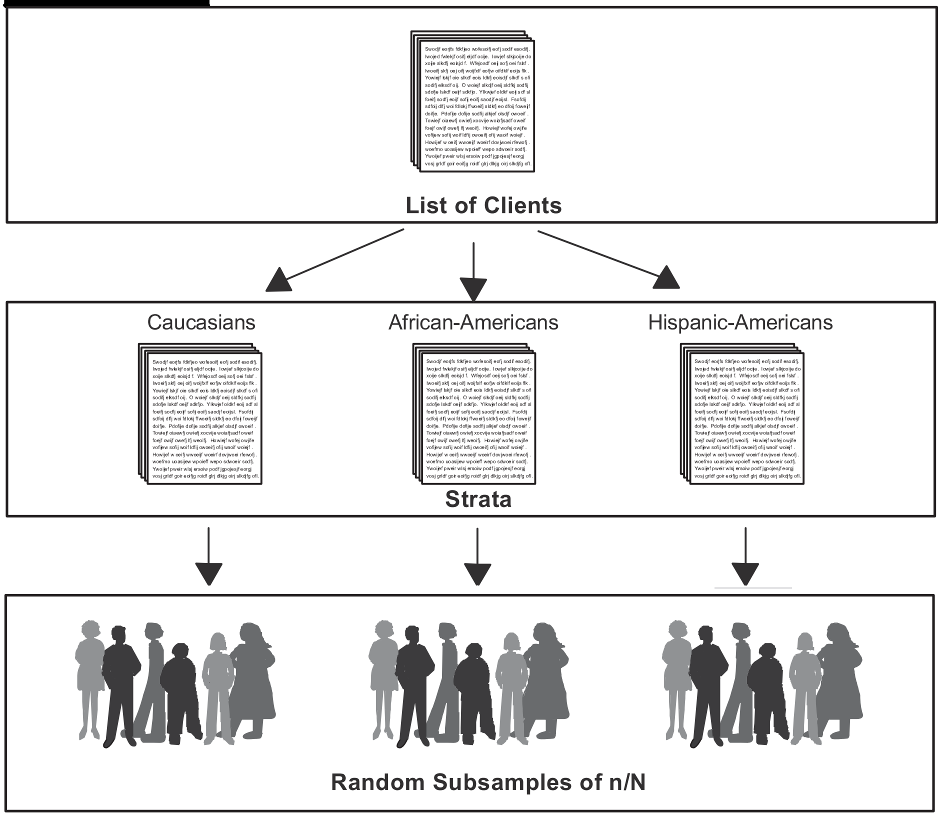
Stratified random sampling, also sometimes called *proportional* or *quota* random sampling, involves dividing your population into homogeneous subgroups and then taking a simple random sample in each subgroup. The following restates this in more formal terms:

> **Objective:** Divide the population into nonoverlapping groups *(strata)* $N_1, N_2, N_3, \ldots N_i$, such that $N_1 + N_2 + N_3 + \ldots + N_i = N$. Then do a simple random sample of $f = n/N$ in each strata.

You might prefer stratified sampling over simple random sampling for several reasons. First, it ensures that you will be able to represent not only the overall population but also key subgroups of the population, especially small minority groups. If you want to be able to talk about subgroups, this may be the only way to ensure effectively you'll be able to do so. If the subgroup is extremely small, you can use different sampling fractions (f) within the different strata to randomly oversample the small group. (Although you'll then have to weight the within-group estimates using the sampling fraction whenever you want overall population estimates.) When you use the same sampling fraction within strata you are conducting *proportionate* stratified random sampling. Using different sampling fractions in the strata is called *disproportionate* stratified random sampling. Second, stratified random sampling has more statistical precision than simple random sampling if the strata or groups are homogeneous. If they are, you should expect the variability within groups to be lower than the variability for the population as a whole. Stratified sampling capitalizes on that fact.

For example, let's say that the population of clients for your agency can be divided as shown in Figure 2–9 into three groups: Caucasian, African American, and Hispanic American. Furthermore, let's assume that both the African Americans and Hispanic Americans are relatively small minorities of the clientele (10 percent and 5 percent, respectively). If you just did a simple random sample of $n = 100$ with a sampling fraction of 10 percent, you would expect by chance alone to get 10 and 5 persons from each of the two smaller groups. And, by chance, you could get even fewer

FIGURE 2–9 **Stratified random sampling**



than that! If you stratify, you can do better. First, you would determine how many people you want to have in each group. Let's say you still want to take a sample of 100 from the population of 1000 clients over the past year; but suppose you think that to say anything about subgroups, you will need at least 25 cases in each group. So, you sample 50 Caucasians, 25 African Americans, and 25 Hispanic Americans. You know that 10 percent of the population, or 100 clients, are African American. If you randomly sample 25 of these, you have a within-stratum sampling fraction of $25/100 = 25\%$. Similarly, you know that 5 percent, or 50 clients, are Hispanic American. So your within-stratum sampling fraction will be $25/50 = 50\%$. Finally, by subtraction you know there are 850 Caucasian clients. Your within-stratum sampling fraction for them is $50/850 =$ about 5.88%. Because the groups are more homogeneous within group than across the population as a whole, you can expect greater statistical precision (less variance), and, because you stratified, you know you will have enough cases from each group to make meaningful subgroup inferences.

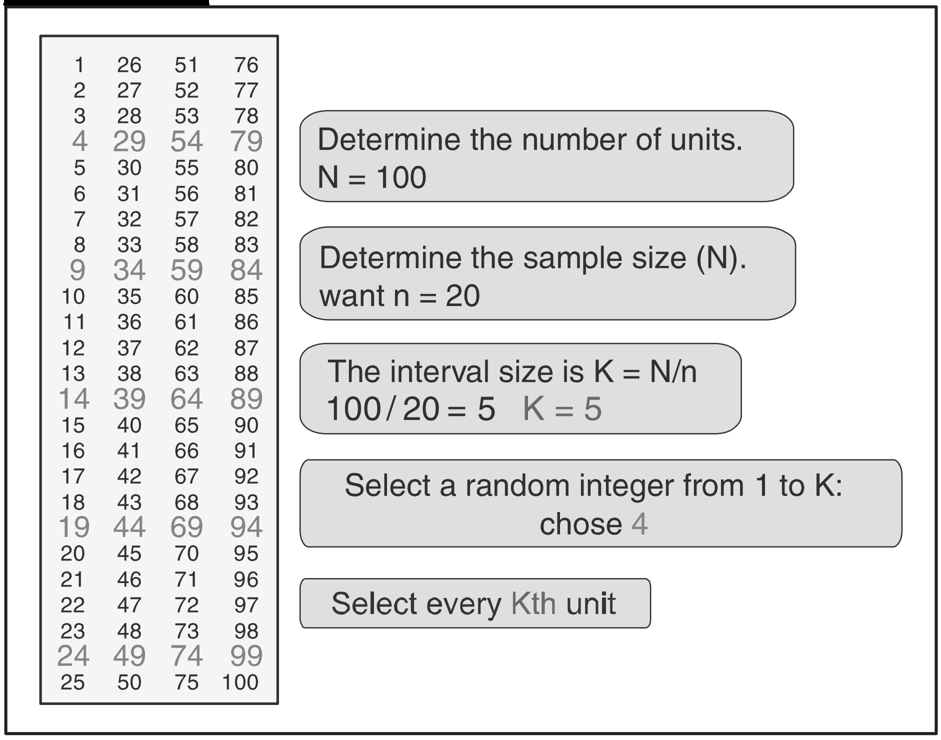
2-4d Systematic Random Sampling

Systematic random sampling is a sampling method where you determine randomly where you want to start selecting in the sampling frame and then follow a rule to select every *x*th element in the sampling frame list (where the ordering of the list is assumed to be random). To achieve a systematic random sample, follow these steps:

systematic random sampling
A sampling method in which you determine randomly where you want to start selecting in the sampling frame and then follow a rule to select every *x*th element the sampling frame list (where the ordering of the list is assumed to be random).

1. Number the units in the population from 1 to N.
2. Decide on the n (sample size) that you want or need.
3. Calculate $k = N/n =$ the interval size.
4. Randomly select an integer between 1 and k.
5. Take every kth unit.

FIGURE 2–10 **Systematic random sampling**



All of this will be much clearer with an example. Let's assume, as shown in Figure 2–10, that you have a population that only has $N = 100$ people in it and that you want to take a sample of $n = 20$. To use systematic sampling, the population must be listed in a random order. The sampling fraction would be $f = 20/100 = 20\%$. In this case, the interval size, k, is equal to $N/n = 100/20 = 5$. Now, select a random integer from 1 to 5. In this example, imagine that you chose 4. Now, to select the sample, start with the 4th unit in the list and take every kth unit (every 5th, because $k = 5$). You would be sampling units 4, 9, 14, 19, and so on, to 100 and you would wind up with 20 units in your sample. For this to work, it is essential that the units in the population be randomly ordered, at least with respect to the characteristics you are measuring. Why would you ever want to use systematic random sampling? For one thing, it is fairly easy to do. You only have to select a single random number to start things off.

It may also be more precise than simple random sampling. Finally, in some situations there is simply no easier way to do random sampling. For instance, I once had to do a study that involved sampling a collection of books in the library. Once selected, I would have to go to the shelf, locate the book, and record when it last circulated. I knew that I had a fairly good sampling frame in the form of the shelf list (which is a card catalog where the entries are arranged in the order they occur on the shelf). To do a simple random sample, I could have estimated the total number of books and generated random numbers to draw the sample; but how would I find book #74,329 easily if that is the number I selected? I couldn't very well count the cards until I came to 74,329! Stratifying wouldn't solve that problem either. For instance, I could have stratified by card catalog drawer and drawn a simple random sample within each drawer. But I'd still be stuck counting cards. Instead, I did a systematic random sample. I estimated the number of books in the entire collection. Let's imagine it was 100,000. I decided that I wanted to take a sample of 1000 for a sampling fraction of $1000/100{,}000 = 1\%$. To get the sampling interval k, I divided $N/n = 100{,}000/100 = 1000$. Then I selected a random integer between 1 and 1000. Let's say I got 257. Next I did a little side study to determine how thick a thousand cards are in the card catalog (taking into account the varying ages of the cards).

Let's say that on average I found that two cards that were separated by 1000 cards were about 2.75 inches apart in the catalog drawer. That information gave me everything I needed to draw the sample. I counted to the 257th by hand and recorded the book information. Then, I took a compass. (Remember those from your high-school math class? They're the funny little metal instruments with a sharp pin on one end and a pencil on the other that you used to draw circles in geometry class.) Then I set the compass at 2.75", stuck the pin end in at the 257th card and pointed with the pencil end to the next card (approximately 1000 books away). In this way, I approximated selecting the 257th, 1257th, 2257th, and so on. I was able to accomplish the entire selection procedure in very little time using this systematic random sampling approach. I'd probably still be there counting cards if I'd tried another random sampling method. (Okay, so I have no life. I got compensated nicely, I don't mind saying, for coming up with this scheme.)

2-4e Cluster (Area) Random Sampling

The problem with random sampling methods when you have to sample a population that's dispersed across a wide geographic region is that you will have to cover a lot of ground geographically to get to each of the units you sampled. Imagine taking a simple random sample of all the residents of New York State to conduct personal interviews. By the luck of the draw, you will wind up with respondents who come from all over the state. Your interviewers are going to have a lot of traveling to do. It is precisely to address this problem that **cluster or area random sampling** was invented.

cluster or area random sampling
A sampling method that involves dividing the population into groups called *clusters*, randomly selecting clusters, and then sampling each element in the selected clusters. This method is useful when sampling a population that is spread across a wide geographic area.

In cluster sampling, you follow these steps:

1. Divide population into clusters (usually along geographic boundaries).
2. Randomly sample clusters.
3. Measure *all* units within sampled clusters.

For instance, Figure 2–11 shows a map of the counties in New York State. Let's say that you have to do a survey of town governments that requires you to go to the towns personally to interview key town officials. If you do a simple random sample of towns statewide, your sample is likely to come from all over the state and you will have to be prepared to cover the entire state geographically. Instead, you can do a cluster sampling of counties, let's say five counties in this example (shaded in the figure). Once these are selected, you go to *every* town government in the five county areas. Clearly this strategy will help you economize on mileage. Instead of having to travel all over the state, you can concentrate exclusively within the counties you selected. Cluster or area sampling is useful in situations like this, and is done primarily for efficiency of administration.

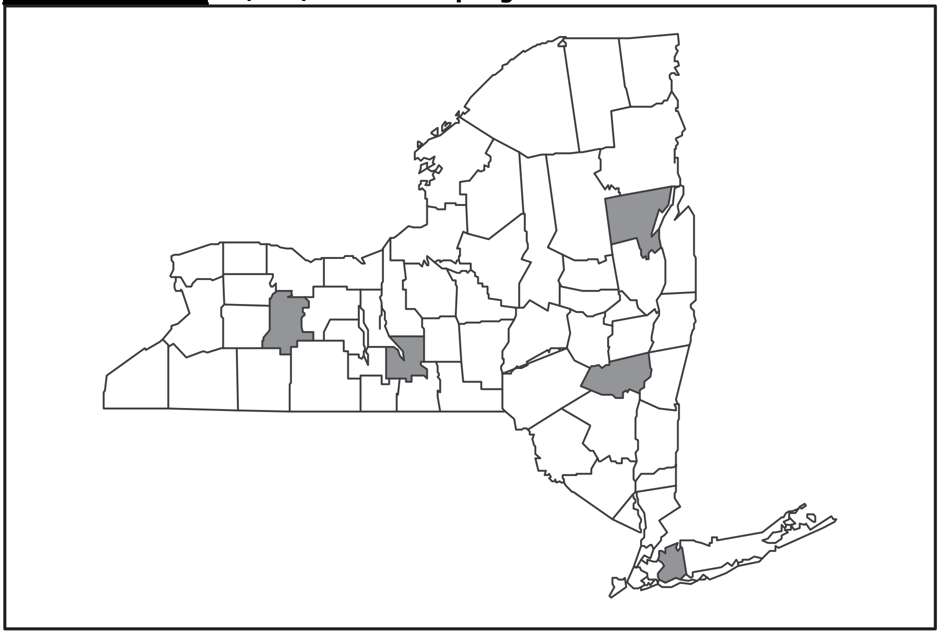
2-4f Multi-Stage Sampling

The four methods covered so far—simple, stratified, systematic, and cluster—are the simplest random sampling strategies. In most real applied social research, you would use sampling methods that are considerably more complex than these simple variations. The most important principle here is that you can combine these simple methods in a variety of useful ways to help you address your sampling needs in the most efficient and effective manner possible. Combining sampling methods is called **multistage sampling**.

multistage sampling
The combining of several sampling techniques to create a more efficient or effective sample than the use of any one sampling type can achieve on its own.

For example, consider the idea of sampling New York State residents for face-to-face interviews. Clearly you would want to do some type of *cluster sampling* as the first stage of the process. You might sample townships or census tracts throughout the state. In cluster sampling, you would then measure everyone in the clusters you selected. Even if you are sampling census tracts, you may not be able to measure *everyone* who is in the census tract. So, you might set up a systematic random sampling

FIGURE 2–11 A county-level map of New York State used for cluster (area) random sampling



process within the clusters. In this case, you would have a two-stage sampling process with stratified samples within cluster samples. Alternatively, consider the problem of sampling students in grade schools. You might begin with a national sample of school districts stratified by economics and educational level. Within selected districts, you might do a simple random sample of schools; within schools, you might do a simple random sample of classes or grades; and, within classes, you might even do a simple random sample of students. In this case, you have three or four stages in the sampling process and you use both stratified and simple random sampling. By combining different sampling methods, you can achieve a rich variety of probabilistic sampling methods to fit a wide range of social research contexts.

2-5 Nonprobability Sampling

nonprobability sampling
Sampling that does not involve random selection.

The difference between nonprobability and probability sampling is that **nonprobability sampling** does not involve random selection and probability sampling does. Does that mean that nonprobability samples aren't representative of the population? Not necessarily; but it does mean nonprobability samples cannot depend upon the rationale of probability theory. At least with a probabilistic sample, you know the odds or probability that you have represented the population well. You can estimate confidence intervals for the statistic. With nonprobability samples, you may or may not represent the population well, and it will often be hard for you to know how well you've done so. In general, researchers prefer probabilistic or random sampling methods over nonprobabilistic ones and consider them to be more accurate and rigorous. However, in some circumstances in applied social research, it is not feasible, practical, or theoretically sensible to use random sampling. In the following paragraphs, I will present a variety of nonprobabilistic sampling alternatives to the probabilistic methods described earlier.

Nonprobability sampling methods are divided into two broad types: *accidental or purposive*. Most sampling methods are purposive in nature because the sampling problem is usually approached with a specific plan in mind. The most important distinctions among these types of sampling methods are between the different types of purposive sampling approaches.

2-5a Accidental, Haphazard, or Convenience Sampling

One of the most common methods of sampling goes under the various titles listed here: *accidental, haphazard*, or *convenience*. I would include in this category the traditional person-on-the-street interviews conducted frequently by television news programs to get a quick (although nonrepresentative) reading of public opinion. I would also argue that the typical use of college students in much psychological research is primarily a matter of convenience. (You don't really believe that psychologists use college students because they think they're representative of the population at large, do you?) In clinical practice, you might use clients available to you as your sample. In many research contexts, you sample by asking for volunteers. Clearly, the problem with all these types of samples is that you have no evidence that they are representative of the populations you're interested in generalizing to, and in many cases, you would suspect that they are not.

2-5b Purposive Sampling

In purposive sampling, you sample with a *purpose* in mind. Usually you would be seeking one or more specific predefined groups. For instance, have you ever run into people in a mall or on the street carrying clipboards and stopping various people and asking to interview them? Most likely, they are conducting a purposive sample (and most likely they are engaged in market research). They might be looking for Caucasian females between 30 and 40 years old. They size up the people passing by and stop people who look to be in that category and ask whether they will participate. One of the first things they're likely to do is verify that the respondent does in fact meet the criteria for being in the sample. Purposive sampling can be useful in situations where you need to reach a targeted sample quickly and where sampling for proportionality is not the primary concern. With a purposive sample, you are likely to get the opinions of your target population, but you are also likely to overweight subgroups in your population that are more readily accessible.

All of the methods that follow can be considered subcategories of purposive sampling methods. You might sample for specific groups or types of people as in modal instance, expert, or quota sampling. You might sample for diversity as in **heterogeneity sampling**, or you might capitalize on informal social networks to identify specific respondents who are hard to locate otherwise, as in **snowball sampling**. In all of these methods, you know what you want—you are sampling with a purpose.

heterogeneity sampling
Sampling for diversity or variety.

Modal Instance Sampling In statistics, the *mode* is the most frequently occurring value in a distribution. In sampling, when you do a **modal instance sample**, you are sampling the most frequent case, or the typical case. Many informal public opinion polls, for instance, interview a typical voter. This sampling approach has a number of problems. First, how do you know what the typical or modal case is? You could say that the modal voter is a person of average age, educational level, and income in the population, but, it's not clear that using the averages of these is the fairest (consider the skewed distribution of income, for instance). In addition, how do you know that those three variables—age, education, income—are the ones most relevant for classifying the typical voter? What if religion or ethnicity is an important determinant of voting decisions? Clearly, modal instance sampling is only sensible for informal sampling contexts.

snowball sampling
A sampling method in which you sample participants based upon referral from prior participants.

modal instance sampling
Sampling for the most typical case.

Expert Sampling **Expert sampling** involves the assembling of a sample of persons with known or demonstrable experience and expertise in some area. Often, you convene such a sample under the auspices of a panel of experts. There are actually two reasons you might do expert sampling. First, it is the best way to elicit the views of persons who have specific expertise. In this case, expert sampling is

expert sampling
A sample of people with known or demonstrable experience and expertise in some area.

essentially just a specific subcase of purposive sampling. The other reason you might use expert sampling is to provide evidence for the validity of another sampling approach you've chosen. For instance, let's say you do modal instance sampling and are concerned that the criteria you used for defining the modal instance is subject to criticism. You might convene an expert panel consisting of persons with acknowledged experience and insight into that field or topic and ask them to examine your modal definitions and comment on their appropriateness and validity. The advantage of doing this is that you aren't out on your own trying to defend your decisions; you have some acknowledged experts to back you. The disadvantage is that even the experts can be, and often are, wrong.

quota sampling
Any sampling method in which you sample until you achieve a specific number of sampled units for each subgroup of a population.

proportional quota sampling
A sampling method in which you sample until you achieve a specific number of sampled units for each subgroup of a population, where the proportions in each group are the same.

nonproportional quota sampling
A sampling method in which you sample until you achieve a specific number of sampled units for each subgroup of a population, where the proportions in each group are not the same.

Quota Sampling In **quota sampling**, you select people nonrandomly according to some fixed quota. The two types of quota sampling are proportional and nonproportional. In **proportional quota sampling**, you want to represent the major characteristics of the population by sampling a proportional amount of each. For instance, if you know the population has 40 percent women and 60 percent men and that you want a total sample size of 100, you should continue sampling until you get those percentages and then stop. So, if you already have the 40 women for your sample, but not the 60 men, you would continue to sample men but even if legitimate women respondents come along, you would not sample them because you have already met your quota. The problem here (as in much purposive sampling) is that you have to decide the specific characteristics on which you will base the quota. Will it be by gender, age, education, race, religion, and so on?

Nonproportional quota sampling is less restrictive. In this method, you specify the minimum number of sampled units you want in each category. Here, you're not concerned with having numbers that match the proportions in the population. Instead, you simply want to have enough to ensure that you will be able to talk about even small groups in the population. This method is the nonprobabilistic analogue of stratified random sampling in that it is typically used to ensure that smaller groups are adequately represented in your sample.

Heterogeneity Sampling You sample for heterogeneity when you want to include all opinions or views, and you aren't concerned about representing these views proportionately. Another term for this is sampling for *diversity*. In many brainstorming or nominal group processes (including concept mapping), you would use some form of heterogeneity sampling because your primary interest is in getting a broad spectrum of ideas, not identifying the average or modal instance ones. In effect, what you would like to be sampling is not people, but ideas. You imagine that there is a universe of all possible ideas relevant to some topic and that you want to sample this population, not the population of people who have the ideas. Clearly, to get all of the ideas, and especially the outlier or unusual ones, you have to include a broad and diverse range of participants. Heterogeneity sampling is, in this sense, almost the opposite of modal instance sampling.

Snowball Sampling In snowball sampling, you begin by identifying people who meet the criteria for inclusion in your study. You then ask them to recommend others they know who also meet the criteria. Although this method would hardly lead to representative samples, at times it may be the best method available. Snowball sampling is especially useful when you are trying to reach populations that are inaccessible or hard to find. For instance, if you are studying the homeless, you are not likely to be able to find good lists of homeless people within a specific geographical area. However, if you go to that area and identify one or two, you may find that they know who the other homeless people in their vicinity are and how you can find them.

Summary

So, that's the basics of sampling methods. Quite a few options, aren't there? How about a table to summarize the choices and give you some idea of when they might be appropriate. Table 2–1 shows each sampling method, when it might best be used, and the major advantages and disadvantages of each.

TABLE 2–1 **Summary of sampling methods**

Probability Sampling Method	Use	Advantages	Disadvantages
Simple random sampling	Anytime	Simple to implement; easy to explain to nontechnical audiences	Requires a sample list (sampling frame) to select from
Stratified random sampling	When concerned about underrepresenting smaller subgroups	Allows you to oversample minority groups to ensure enough for subgroup analyses	Requires a sample list (sampling frame) from which to select
Systematic random sampling	When you want to sample every *k*th element in an ordered set	Does not require that you count through all of the elements in the list to find the ones randomly selected	If the order of elements is nonrandom, there could be systematic bias
Cluster (area) random sampling	When organizing geographically makes sense	Is more efficient than other methods when sampling across a geographically dispersed area	Is usually not used alone; is coupled with other methods in a multistage approach
Multistage random sampling	Anytime	Combines sophistication with efficiency	Can be complex and difficult to explain to nontechnical audiences
Non-probability Sampling Methods			
Accidental, haphazard, or convenience nonprobability sampling	Anytime	Is very easy to do; almost like not sampling at all	Has very weak external validity; is likely to be biased
Modal instance purposive nonprobability sampling	When you want to measure only a typical respondent	Is easily understood by nontechnical audiences	Results limited to only the modal case; has little external validity
Modal purposive nonprobability sampling	As an adjunct to other sampling strategies	Experts can provide opinions to support research conclusions	Is likely to be biased; has limited external validity
Quota purposive nonprobability sampling	When you want to represent subgroups	Allows for oversampling smaller subgroups	Is likely to be more biased than stratified random sampling; often depends on who comes along when
Heterogeneity purposive nonprobability sampling	When you want to sample for diversity or variety	Is easy to implement and explain; is useful when you're interested in sampling for variety rather than representativeness	Won't represent population views proportionately
Snowball purposive nonprobability sampling	With hard-to-reach populations	Can be used when there is no sampling frame	Has low external validity

Sampling is a critical component in virtually all social research. While I've presented a wide variety of sampling methods in this chapter, it's important that you keep them in perspective. The key is not which sampling method you use. The key is external validity—how valid the inferences from your sample are. You can have the best sampling method in the world and it won't guarantee that your generalizations are valid (although it does help!). Alternatively, you can use a relatively weak nonprobability sampling method and find that it is perfectly useful for your context. Ultimately whether your generalizations from your study to other persons, places, or times are valid is a judgment. Your critics, readers, friends, supporters, funders, and so on, will judge the quality of your generalizations, and they may not even agree with each other in their judgment. What might be convincing to one person or group may fail with another. Your job as a social researcher is to create a sampling strategy that is appropriate to the context and will ensure that your generalizations are as convincing as possible to as many audiences as is feasible.

Login to the Online Edition of your text at www.atomicdog.com to find additional resources located in the Study Guide at the end of each chapter.