

Chapter 7. Sampling Techniques

Introduction to Sampling

Distinguishing Between a Sample and a Population

Simple Random Sampling

Step 1. Defining the Population

Step 2. Constructing a List

Step 3. Drawing the Sample

Step 4. Contacting Members of the Sample

Stratified Random Sampling

Convenience Sampling

Quota Sampling

Thinking Critically About Everyday Information

Sample Size

Sampling Error

Evaluating Information From Samples

Case Analysis

General Summary

Detailed Summary

Key Terms

Review Questions/Exercises

Introduction to Sampling

The way in which we select a sample of individuals to be research participants is critical. How we select participants (random sampling) will determine the population to which we may generalize our research findings. The procedure that we use for assigning participants to different treatment conditions (random assignment) will determine whether bias exists in our treatment groups (Are the groups equal on all known and unknown factors?). We address random sampling in this chapter; we will address random assignment later in the book.

If we do a poor job at the sampling stage of the research process, the integrity of the entire project is at risk. If we are interested in the effect of TV violence on children, which children are we going to observe? Where do they come from? How many? How will they be selected? These are important questions. Each of the sampling techniques described in this chapter has advantages and disadvantages.

Distinguishing Between a Sample and a Population

Before describing sampling procedures, we need to define a few key terms. The term **population** means *all* members that meet a set of specifications or a specified criterion. For example, the population of the United States is defined as all people residing in the United States. The population of New Orleans means all people living within the city's limits or boundary. A population of inanimate objects can also exist, such as all automobiles manufactured in Michigan in the year 2003. A single member of any given population is referred to as an **element**. When only some elements are selected from a population, we refer to that as a **sample**; when all elements are included, we call it a **census**. Let's look at a few examples that will clarify these terms.

Two research psychologists were concerned about the different kinds of training that graduate students in clinical psychology were receiving. They knew that different programs emphasized different things, but they did not know which clinical orientations were most popular. Therefore, they prepared a list of *all* doctoral programs in clinical psychology (in the United States) and sent each of them a questionnaire regarding aspects of their program. The response to the survey was excellent; nearly 95% of the directors of these programs returned the completed questionnaire. The researchers then began analyzing their data and also classifying schools into different clinical orientations: psychoanalytic, behavioristic, humanistic, Rogerian, and so on. When the task was complete, they reported the percentage of schools having these different orientations and described the orientations that were most popular, which were next, and so on. They also described other aspects of their data. The study was written up and submitted for publication to one of the professional journals dealing with matters of clinical psychology. The editor of the journal read the report and then returned it with a letter rejecting the manuscript for publication. In part, the letter noted that the manuscript was not publishable at this time because the proper

statistical analyses had not been performed. The editor wanted to know whether the differences in orientation found among the different schools were significant or if they were due to chance.

The researchers were unhappy, and rightly so. They wrote back to the editor, pointing out that their findings were not estimates based on a sample. They had surveyed *all* training programs (that is, the population). In other words, they had obtained a census rather than a sample. Therefore, their data were exhaustive; they included all programs and described what existed in the real world. The editor would be correct only if they had sampled some schools and then wanted to generalize to all schools. The researchers were not asking whether a sample represented the population; they were dealing with the population.

A comparable example would be to count *all students* (the population) enrolled in a particular university and then report the number of male and female students. If we found that 60% of the students were female, and 40% male, it would be improper and irrelevant to ask whether this difference in percentage is significantly different from chance. The fact is that the percentages that exist in the school population are parameters. They are not estimates derived from a sample. Had we taken a small sample of students and found this 60/40 split, it would then be appropriate to ask whether differences this large could have occurred by chance alone.

Data derived from a sample are treated *statistically*. Using sample data, we calculate various statistics, such as the mean and standard deviation. These sample statistics summarize (describe) aspects of the sample data. These data, when treated with other statistical procedures, allow us to make certain inferences. From the sample statistics, we make corresponding estimates of the population. Thus, from the sample mean, we estimate the population mean; from the sample standard deviation, we estimate the population standard deviation.

The above examples illustrate a problem that can occur when the terms *population* and *sample* are confused. The accuracy of our estimates depends on the extent to which the sample is representative of the population to which we wish to generalize.

Simple Random Sampling

Researchers use two major sampling techniques: probability sampling and nonprobability sampling. With **probability sampling**, a researcher can specify the probability of an element's (participant's) being included in the sample. With **nonprobability sampling**, there is no way of estimating the probability of an element's being included in a sample. If the researcher's interest is in generalizing the findings derived from the sample to the general population, then probability sampling is far more useful and precise. Unfortunately, it is also much more difficult and expensive than nonprobability sampling.

Probability sampling is also referred to as **random sampling** or representative sampling. The word *random* describes the procedure used to select elements (participants, cars, test items) from a population.

When random sampling is used, each element in the population has an equal chance of being selected (simple random sampling) or a known probability of being selected (stratified random sampling). The sample is referred to as *representative* because the characteristics of a properly drawn sample represent the parent population in all ways.

One caution before we begin our description of simple random sampling: Random *sampling* is different from random *assignment*. Random assignment describes the process of placing participants into different experimental groups. We will discuss random assignment later in the book.

Step 1. Defining the Population

Before a sample is taken, we must first define the population to which we want to generalize our results. The population of interest may differ for each study we undertake. It could be the population of professional football players in the United States or the registered voters in Bowling Green, Ohio. It could also be all college students at a given university, or all sophomores at that institution. It could be female students, or introductory psychology students, or 10-year-old children in a particular school, or members of the local senior citizens center. The point should be clear; the sample should be drawn from the population to which you want to generalize—the population in which you are interested.

It is unfortunate that many researchers fail to make explicit their population of interest. Many investigators use only college students in their samples, yet their interest is in the adult population of the United States. To a large extent, the generalizability of sample data depends on what is being studied and the inferences that are being made. For example, imagine a study that sampled college juniors at a specific university. Findings showed that a specific chemical compound produced pupil dilation. We would not have serious misgivings about generalizing this finding to all college students, even tentatively to all adults, or perhaps even to some nonhuman organisms. The reason for this is that physiological systems are quite similar from one person to another, and often from one species to another. However, if we find that controlled exposure to unfamiliar political philosophies led to radicalization of the experimental participants, we would be far more reluctant to extend this conclusion to the general population.

Step 2. Constructing a List

Before a sample can be chosen randomly, it is necessary to have a complete list of the population from which to select. In some cases, the logistics and expense of constructing a list of the entire population is simply too great, and an alternative procedure is forced upon the investigator. We could avoid this problem by restricting our population of interest—by defining it narrowly. However, doing so might increase the difficulty of finding or constructing a list from which to make our random selection. For example, you would have no difficulty identifying female students at any given university and then constructing a list of their names from which to draw a random sample. It would be more difficult to

identify female students coming from a three-child family, and even more difficult if you narrowed your interest to firstborn females in a three-child family. Moreover, defining a population narrowly also means generalizing results narrowly.

Caution must be exercised in compiling a list or in using one already constructed. The population list from which you intend to sample must be both recent and exhaustive. If not, problems can occur. By an **exhaustive list**, we mean that *all* members of the population must appear on the list. Voter registration lists, telephone directories, homeowner lists, and school directories are sometimes used, but these lists may have limitations. They must be up to date and complete if the samples chosen from them are to be truly representative of the population. In addition, such lists may provide very biased samples for some research questions we ask. For example, a list of homeowners would not be representative of all individuals in a given geographical region because it would exclude transients and renters. On the other hand, a ready-made list is often of better quality and less expensive to obtain than a newly constructed list would be.

Some lists are available from a variety of different sources. Professional organizations, such as the American Psychological Association, the American Medical Association, and the American Dental Association, have directory listings with mailing addresses of members. Keep in mind that these lists do not represent all psychologists, physicians, or dentists. Many individuals do not become members of their professional organizations. Therefore, a generalization would have to be limited to those professionals listed in the directory. In universities and colleges, complete lists of students can be obtained from the registrar.

Let's look at a classic example of poor sampling in the hours prior to a presidential election. Information derived from sampling procedures is often used to predict election outcomes. Individuals in the sample are asked their candidate preferences before the election, and projections are then made regarding the likely winner. More often than not, the polls predict the outcome with considerable accuracy. However, there are notable exceptions, such as the 1936 *Literary Digest* magazine poll that predicted "Landon by a Landslide" over Roosevelt, and predictions in the U.S. presidential election of 1948 that Dewey would defeat Truman.

We have discussed the systematic error of the *Literary Digest* poll. Different reasons resulted in the wrong prediction in the 1948 presidential election between Dewey and Truman. Polls taken in 1948 revealed a large undecided vote. Based partly on this and early returns on the night of the election, the editors of the *Chicago Tribune* printed and distributed their newspaper before the election results were all in. The headline in bold letters indicated that Dewey defeated Truman. Unfortunately for them, they were wrong. Truman won, and the newspaper became a collector's item.

One analysis of why the polls predicted the wrong outcome emphasized the consolidation of opinion for many undecided voters. It was this undecided group that proved the prediction wrong. Pollsters did

not anticipate that those who were undecided would vote in large numbers for Truman. Other factors generally operate to reduce the accuracy of political polls. One is that individuals do not always vote the way they say they are going to. Others may intend to do so but change their mind in the voting booth. Also, the proportion of potential voters who actually cast ballot differs depending upon the political party and often upon the candidates who are running. Some political analysts believe (along with politicians) that even the position of the candidate's name on the ballot can affect the outcome (the debate regarding butterfly ballots in Florida during the 2000 presidential election comes to mind).

We will describe the mechanics of random sampling shortly, but we want to note again that in some cases random sampling procedures simply are not possible. This is the case for very large populations. Because random sampling requires a listing of all members of a population, the larger the population the more difficult it becomes.

Step 3. Drawing the Sample

After a list of population members has been constructed, various random sampling options are available. Some common ones include tossing dice, flipping coins, spinning wheels, drawing names out of a rotating drum, using a table of random numbers, and using computer programs. Except for the last two methods, most of the techniques are slow and cumbersome. Tables of random numbers are easy to use, accessible, and truly random. Here is a website that provides a random number table, as well as a way to generate random numbers ([website](#)).

Let's look at the procedures for using the table. The first step is to assign a number to each individual on the list. If there were 1,000 people in the population, you would number them 0 to 999 and then enter the table of random numbers. Let us assume your sample size will be 100. Starting anywhere in the table, move in any direction you choose, preferably up and down. Since there are 1,000 people on your list (0 through 999) you must give each an equal chance of being selected. To do this, you use three columns of digits from the tables. If the first three-digit number in the table is 218, participant number 218 on the population list is chosen for the sample. If the next three-digit number is 007, the participant assigned number 007 (or 7) is selected. Continue until you have selected all 100 participants for the sample. If the same number comes up more than once, it is simply discarded.

In the preceding fictional population list, the first digit (9) in the total population of 1,000 (0-999) was large. Sometimes the first digit in the population total is small, as with a list of 200 or 2,000. When this happens, many of the random numbers encountered in the table will not be usable and therefore must be passed up. This is very common and does not constitute a sampling problem. Also, tables of random numbers come in different column groupings. Some come in columns of two digits, some three, some four, and so on. These differences have no bearing on randomness. Finally, it is imperative that you not violate the random selection procedure. Once the list has been compiled and the process of selection has

begun, the table of random numbers dictates who will be selected. The experimenter should not alter this procedure.

A more recent method of random sampling uses the special functions of computer software. Many population lists are now available as software databases (such as Excel, Quattro Pro, Lotus123) or can be imported to such a database. Many of these database programs have a function for generating a series of random numbers and a function for selecting a random sample from a range of entries in the database. We also mentioned above that numerous internet sites can generate random numbers. After you learn the particular menu selections to perform these tasks, these methods of random sampling are often the simplest.

Step 4. Contacting Members of a Sample

Researchers using random sampling procedures must be prepared to encounter difficulties at several points. As we noted, the starting point is an accurate statement that identifies the population to which we want to generalize. Then we must obtain a listing of the population, accurate and up-to-date, from which to draw our sample. Further, we must decide on the random selection procedure that we wish to use. Finally, we must contact each of those selected for our sample and obtain the information needed. Failing to contact all individuals in the sample can be a problem, and the representativeness of the sample can be lost at this point.

To illustrate what we mean, assume that we are interested in the attitudes of college students at your university. We have a comprehensive list of students and randomly select 100 of them for our sample. We send a survey to the 100 students, but only 80 students return it. We are faced with a dilemma. Is the sample of 80 students who participated representative? Because 20% of our sample was not located, does our sample underrepresent some views? Does it overrepresent other views? In short, can we generalize from our sample to the college population? Ideally, all individuals in a sample should be contacted. As the number contacted decreases, the risk of bias and not being representative increases.

Thus, in our illustration, to generalize to the college population would be to invite risk. Yet we do have data on 80% of our sample. Is it of any value? Other than simply dropping the project or starting a new one, we can consider an alternative that other researchers have used. In preparing our report, we would first clearly acknowledge that not all members of the sample participated and therefore the sample may not be random—that is, representative of the population. Then we would make available to the reader or listener of our report the number of participants initially selected and the final number contacted, the number of participants cooperating, and the number not cooperating. We would attempt to assess the reason or reasons participants could not be contacted and whether differences existed between those for whom there were data and those for whom there were no data. If no obvious differences were found, we could feel a little better about the sample's being representative. However, if any pattern of

differences emerged, such as sex, education, or religious beliefs, a judgment would have to be made regarding how seriously the differences could have affected the representativeness of the sample.

Differences on any characteristic between those who participated and those who did not should not automatically suggest that the information they might give would also differ. Individuals can share many common values and beliefs, even though they may differ on characteristics such as sex or education. In situations requiring judgments, such as the one described, the important thing is for the researcher to describe the strengths and weaknesses of the study (especially telling the reader that only 80 of the 100 surveys were returned), along with what might be expected as a result of them. Alert the reader or listener to be cautious in interpreting the data, and provide them with the information necessary to make an informed judgment.

The problem just described may be especially troublesome when surveys or questionnaires deal with matters of a personal nature. Individuals are usually reluctant to provide information on personal matters, such as sexual practices, religious beliefs, or political philosophy. The more personal the question, the fewer the number of people who will respond. With surveys or questionnaires of this nature, a large number of individuals may refuse to cooperate or refuse to provide certain information. Some of these surveys have had return rates as low as 20%. If you are wondering what value publishing such data has when derived from such a low return rate, you are in agreement with us. We, too, wonder why such data are published. Even if we knew the population from which the sample was drawn and if the sample was randomly selected, a return rate as low as 20% is virtually useless in terms of generalizing findings from the sample to the population. Those individuals responding to a survey (20% of the sample) could be radically different from the majority of individuals not responding (80% of the sample).

Let's apply these four steps of random sampling to our TV violence study. Our first step is to define the population. We might begin by considering the population as all children in the United States that are 5–15 years old. Our next step will be to obtain an exhaustive list of these children. Using U. S. Census data would be one approach, although the task would be challenging and the Census does miss many people. The third step is to select a random sample. As noted earlier in the chapter, the simplest technique would be to use a database of the population and instruct the database software to randomly select children from the population. The number to be selected is determined by the researcher and is typically based on the largest number that can be sampled given the logistical resources of the researcher. Of course, the larger the sample, the more accurately it will represent the population. In fact, formulas can be used to determine sample size based on the size of the population, the amount of variability in the population, the estimated size of the effect, and the amount of sampling error that the researcher decides is acceptable (refer to statistics books for specifics). After the sample is selected from the population, the final step is to contact the parents of these children to obtain consent to participate. You will need to make phone calls and send letters. Again, this will be a challenge; you expect that you will be unable to contact

a certain percentage, and that a certain percentage will decline to participate. All this effort, and we have not even begun to talk about collecting data from these children.

From this example, it is clear that random sampling can require an incredible amount of financial resources. As noted earlier in the chapter, we have two options. We can define the population more narrowly (perhaps the 5- to 15-year-olds in a particular school district) and conduct random sampling from this population, or we can turn to a sampling technique other than probability sampling. Before we discuss these nonprobability sampling techniques, let's look at one other form of probability sampling.

Stratified Random Sampling

This procedure known as **stratified random sampling** is also a form of probability sampling. To stratify means to classify or to separate people into groups according to some characteristics, such as position, rank, income, education, sex, or ethnic background. These separate groupings are referred to as *subsets* or *subgroups*. For a stratified random sample, the population is divided into groups or strata. A random sample is selected from each stratum based upon the percentage that each subgroup represents in the population. Stratified random samples are generally more accurate in representing the population than are simple random samples. They also require more effort, and there is a practical limit to the number of strata used. Because participants are to be chosen randomly from each stratum, a complete list of the population within each stratum must be constructed. Stratified sampling is generally used in two different ways. In one, primary interest is in the representativeness of the sample for purposes of commenting on the population. In the other, the focus of interest is comparison between and among the strata.

Let's look first at an example in which the population is of primary interest. Suppose we are interested in the attitudes and opinions of university faculty in a certain state toward faculty unionization. Historically, this issue has been a very controversial one evoking strong emotions on both sides. Assume that there are eight universities in the state, each with a different faculty size (faculty size = $500 + 800 + 900 + 1,000 + 1,400 + 1,600 + 1,800 + 2,000 = 10,000$). We could simply take a simple random sample of all 10,000 faculty and send those in the sample a carefully constructed attitude survey concerning unionization. After considering this strategy, we decide against it. Our thought is that universities of different size may have marked differences in their attitudes, and we want to be sure that each university will be represented in the sample in proportion to its representation in the total university population. We know that, on occasion, a simple random sample will not do this. For example, if unionization is a particularly "hot" issue on one campus, we may obtain a disproportionate number of replies from that faculty. Therefore, we would construct a list of the entire faculty for each university and then sample randomly within each university in proportion to its representation in the total faculty of 10,000. For example, the university with 500 faculty members would represent 5% of our sample; assuming a total sample size of 1,000, we would randomly select 50 faculty from this university. The university with 2,000

faculty would represent 20% of our sample; thus, 200 of its faculty would be randomly selected. We would continue until our sample was complete. It would be possible but more costly and time consuming to include other strata of interest—for example, full, associate, and assistant professors. In each case, the faculty in each stratum would be randomly selected.

As previously noted, stratified samples are sometimes used to optimize group comparisons. In this case, we are not concerned about representing the total population. Instead, our focus is on comparisons involving two or more strata. If the groups involved in our comparisons are equally represented in the population, a single random sample could be used. When this is not the case, a different procedure is necessary. For example, if we were interested in making comparisons between whites and blacks, a simple random sample of 100 people might include about 85 to 90 whites and only 10 to 15 blacks. This is hardly a satisfactory sample for making comparisons. With a stratified random sample, we could randomly choose 50 whites and 50 blacks and thus optimize our comparison. Whenever strata rather than the population are our primary interest, we can sample in different proportions from each stratum.

Although random sampling is optimal from a methodological point of view, it is not always possible from a practical point of view. Let's look at the advantages and disadvantages of several other sampling techniques.

Convenience Sampling

Convenience sampling is used because it is quick, inexpensive, and convenient. Convenience samples are useful for certain purposes, and they require very little planning. Researchers simply use participants who are available at the moment. The procedure is casual and easy, relative to random sampling. Contrast using any available participants with random sampling, where you must (1) have a well-defined population, (2) construct a list of members of the population if one is not available, (3) sample randomly from the list, and (4) contact and use as many individuals from the list as possible. Convenience sampling requires far less effort. However, such convenience comes with potential problems, which we will describe. Convenience samples are nonprobability samples. Therefore, it is not possible to specify the probability of any population element's being selected for the sample. Indeed, it is not possible to specify the population from which the sample was drawn.

A number of examples of convenience sampling can be given. In shopping malls or airports, individuals are selected as they pass a certain location and interviewed concerning issues, candidates, or other matters. Phone surveys may be based on anyone answering the phone between the hours of 9 A.M. and 5 P.M. Politicians use convenience sampling to determine the attitudes of those they represent when they report on the number of letters voluntarily sent to them by their constituents. Statements such as the following are quite common: "My mail is running about 4 to 1 in favor of House Bill 865A. I guess I

know how my constituents feel about the issue.” Unfortunately, many of these samples are virtually without merit. We do not know what population (whom) they represent.

These and other examples can be used to illustrate the point. Observations at airports may overrepresent high-income groups, whereas observations taken at bus terminals may overrepresent low-income groups. Surveys taken at a rock concert would likely be different from those taken at a symphony concert. In the case of political attitudes, we do know that many special interest groups make it a matter of policy to write letters to their political representatives. A thousand people vitally concerned about an issue may write more letters than a million people who are indifferent. Polls taken on the Internet have become more popular and suffer from the same drawbacks. The point we are making is this: Because the population from which the sample came is unknown, it is unclear to whom the data can be generalized. We can generalize to known populations, but only with some risk. We will have more to say about this below.

The examples used here are extreme and the problems obvious, but there are instances where these problems are not as serious or as apparent. In these instances, some researchers believe that convenience sampling is a good alternative to random sampling.

As noted earlier, most laboratory research in psychology, human and nonhuman, uses a convenience sampling procedure. Some universities require that students taking the introductory psychology course serve as participants in research projects of their choosing. When participants are required to participate in research and are allowed to choose certain experiments over others, then for any given experiment, it is simply impossible to specify the population to which the sample data can be generalized. In other words, to what individuals, other than those of the sample, are the data relevant? We have here a sample in search of a population.

Requiring students to participate in research serves several purposes. It assures that each student has an opportunity to learn firsthand about scientific research. In this regard, an attempt is made to make participation in research a worthwhile educational experience. It also assures that participants are generally available for research, thus serving the purposes of the researcher and that of psychology as a science. The system requiring participants to participate in research of their choosing operates in the following manner: Research projects to be undertaken are listed on a bulletin board (or a Web site), with a brief description of the project and a sign-up sheet indicating the time, place, and experimenter. If our earlier description was clear, you will recognize this as a convenience sampling procedure. Although the students are required to participate in research, they choose the particular project in which to participate. If students are available at a given time, and the particular experiment appeals to them, they simply sign their name on the sign-up sheet.

Frequently, the description on the sign-up sheet is neutral, but sometimes it is not. The titles alone—for example, Reaction Time to Electrical Stimulation, Problem Solving and Cognitive Skills, or

Personality Assessment—are often threatening to some individuals. Obviously, these are not neutral topics, and you can anticipate what may occur. Participants concerned about the words “electrical stimulation” will avoid that particular experiment. Those concerned about their problem-solving ability may think they are to be evaluated and thus avoid that particular experiment. And so it goes. Although all students may participate in research, certain experiments may attract students with certain characteristics. In principle, students with different characteristics represent different populations. Even experiments with titles and descriptions that appear neutral may attract certain kinds of participants over other kinds. We will restate the point we stated earlier: Students participating in these experiments may be thought of as a sample of students from a population of students with certain characteristics, but a population that we cannot identify. Again, we ask: “To what individuals, other than those in the sample, are the data relevant or generalizable?” More concretely, conclusions drawn from the data of students who signed up for a study using electrical stimulation could be very different than if the data had come from students who avoided the experiment. It would be improper to generalize the findings to all students.

Some researchers using convenience samples are not concerned about the population to which they can properly generalize because their interest is in assessing the relationship between the independent and dependent variables. Their concern is focused on internal validity (minimizing confounding) rather than on external validity (generalizing their findings). Others, however, interested in generalizing from the sample to the population represented by it, argue that there is no good reason for assuming that students making up convenience samples are different from the general population of college students. Therefore, they would be willing to generalize their findings to all college students.

A similar argument is made by researchers using convenience samples of nonhuman participants, such as rats, cats, and dogs. In this case, sample findings are generalized to all rats, cats, or dogs of a given strain. The argument that the sample results are generalizable to all college students or to all animals of a given species and strain may be correct, but the argument is not based on firm theoretical grounds, nor can it seek support from statistical sampling theory. The argument is based more on faith and intuition than on objective argument derived from sampling theory. To what populations can convenience samples be generalized? The population to which it is permissible to generalize is that from which the sample was drawn. Strictly speaking, the population from which the sample was drawn is unknown. Because the sample was not drawn randomly from a list of some well-defined population, the population to which the sample findings can be generalized cannot be identified. A real dilemma exists. We have a sample in search of a population. We want to generalize our results beyond the sample, but to whom? This dilemma is inevitable when convenience sampling is used. Under these circumstances, statements concerning generality should be cautious, conservative, and appropriately qualified.

Had a listing of all introductory psychology students at a given university been available, and an adequate number of participants selected *randomly* from the list for any given experiment, we would not

face the dilemma of generalizing our results. If we randomly drew names from this list, our sample would represent the population of introductory psychology students at the university. However, generalizing from our sample to introductory students at other universities would entail some risk. Our sample may not be representative of the population of introductory psychology students at other universities.

Quota Sampling

In many large-scale applications of sampling procedures, it is not always possible or desirable to list all members of the population and randomly select elements from that list. The reasons for using any alternative procedures include cost, timeliness, and convenience. One alternative procedure is **quota sampling**.

This technique is often used by market researchers and those taking political polls. Usually, when this technique is used, the population of interest is large and there are no ready-made lists of names available from which to sample randomly. The Gallup Poll is one of the best known and well conducted polls to use quota sampling. This poll frequently reports on major public issues and on presidential elections. The results of the poll are syndicated for a fee that supports it. In this quota sampling procedure, localities are selected and interviewers are assigned a starting point, a specified direction, and a goal of trying to meet quotas for subsets (ethnic origins, political affiliations, and so on) selected from the population. Although some notable exceptions have occurred, predictions of national elections over the past few years have been relatively accurate—certainly, much more so than guesswork.

With the quota sampling procedure, we first decide which subgroups of the population interest us. This, in turn, is dictated by the nature of the problem being investigated (the question being asked). For issues of national interest (such as abortion, drug use, or political preference), frequently used subsets are age, race, sex, socioeconomic level, and religion. The intent is to select a sample whose frequency distribution of characteristics reflects that of the population of interest. Obviously, it is necessary to know the percentage of individuals making up each subset of the population if we are to match these percentages in the sample. For example, if you were interested in ethnic groups such as Italians, Germans, Russians, and so on, and knew their population percentages, you would select your sample so as to obtain these percentages.

Within each subset, participants are *not* chosen randomly. This is simply because there are usually no ready-made lists from which the researcher can select randomly. Often individuals are selected in the sample on the basis of availability. For this reason, quota sampling is less expensive. It would not be so if lists of the population of interest had to be constructed. However, if exhaustive ready-made lists were conveniently available for the population of interest, then choosing participants randomly would be possible and preferable. In the absence of such lists, it is much more convenient to select quotas by knocking on doors, telephoning numbers, or sending mailings until the sample percentages for subsets

match those of the population. Obviously, even though the quotas may be achieved and the sample may match the population percentages in terms of subsets, the sample may still not represent (reflect) the population to which we wish to generalize.

Often interviewers, for sampling purposes, concentrate on areas where large numbers of people are likely to be. This could bias the findings. As we noted earlier, samples taken in airports may overrepresent high-income groups, whereas those at a bus or rail depots may overrepresent low-income groups. Samples at either place may underrepresent those who seldom travel. Also, people who are home during the day, and are therefore available for house-to-house interviews or telephone calls, may be quite different in important ways from those who are not home. In this respect, quota sampling and convenience sampling are similar. In spite of these difficulties, the quota system is widely used and will unquestionably continue to be so for economic and logistic reasons.

We have just reviewed four sampling techniques: simple random sampling, stratified random sampling, convenience sampling, and quota sampling. Table 7.1 presents one more example of each technique as it relates to our TV violence study, noting the advantages and limitations of each. For the purposes of these examples, we will define the population as all middle-school children in the Conway, Arkansas, school district (approximately 1000 students).

Table 7.1 Examples, Advantages, and Limitations of the Four Sampling Techniques			
SAMPLING TECHNIQUE	EXAMPLE	ADVANTAGES	LIMITATIONS
Simple random sampling	The names of all 1,000 children are placed into a computer database. The computer is then instructed to randomly select 100 names. These children and their parents are then contacted.	Representative of the population	May be difficult to obtain the list May be more expensive
Stratified random sampling	The names of all 1,000 children are placed into a computer database and organized by grade (sixth, seventh, eighth). The computer is then instructed to randomly select 35 names from each of the three grades. These children and their parents are then contacted.	Representative of the population	May be difficult to obtain the list May be more expensive
Convenience sampling	The researcher knows one of the middle-school teachers, and the teacher volunteers her 35 students for the study. These children and their parents are then contacted.	Simple Easy Convenient No complete member list needed	May not be representative of the population
Quota sampling	Using the middle-school directory, the researcher selects the first 20 sixth-grade boys, the first 20 sixth-grade girls, the first 20 seventh-grade boys, the first 20 seventh-grade girls, the first 20 eighth-grade boys, and the first 20 eighth-grade girls. These children and their parents are then contacted.	Simple Easy Convenient No complete member list needed	May not be representative of the population

As we have seen, the method of sampling is an important issue in research. See “Thinking Critically About Everyday Information” for a news report in which the method of sampling may be problematic.

Thinking Critically About Everyday Information: Can Cheating Make Your Marriage Stronger?

The following report of a study in Italy was obtained from the online edition of *Weekly World News*. The title of the report is “New Study Reveals...Cheating Makes Your Marriage Stronger.” A portion of the report reads:

I started the analysis project to discover how damaging infidelity was to marriages,” says Dr. Ostertag. “I was as surprised as everyone when the numbers proved that cheating on your spouse is actually good for your marriage.” According to the scientific survey, the more extramarital flings a couple enjoys, the more likely they are to remain together and the happier they will be. “Some of the strongest unions I studied included spouses who each were involved in repeated extramarital affairs throughout the relationship,” explains Dr. Ostertag. “My findings have turned our preconceived notion of the strength of monogamy on its head.”

Carefully consider the following questions:

- Although the article does not provide methodological details (a shortcoming of many media reports), of the sampling techniques discussed in this chapter, which comes closest to the type that was probably used for this study?
- Think about how the researcher might have sampled married couples. How could biased sampling explain the pattern of results?
- What would be the most effective (accurate) method of sampling in order to answer the research question?

SOURCE: Retrieved June 10, 2003, online at

http://www.weeklyworldnews.com/features/revelations_story.cfm?instanceid=51782

Sample Size

We briefly mentioned the issue of sample size earlier when we discussed random sampling for our TV violence study. We are always confronted with the question of how large a sample should be drawn. The size of the sample depends on various considerations, including population variability, statistical issues, economic factors, availability of participants, and the importance of the problem.

In inferential statistics, the sample size required depends on how big a difference between two groups you want to be able to detect. With large sample sizes, small differences can be detected. If the sample size is sufficiently large, virtually any population difference will result in statistical significance. On the other hand, the smaller the sample size, the larger must be the population differences to achieve statistical significance. Stated another way, other things being equal, the greater the sample size, the less is the probability of drawing a conclusion that is in error. In carefully conducted survey research, the sample size determines how closely the sample values approximate the population values. Assuming valid sampling procedures, the larger the sample, the more closely (on average) will the sample values approximate the population values. However, the relationship between sample size and sensitivity is a curve of diminishing returns. Beyond a certain point, the cost and effort required to achieve greater sensitivity becomes disproportionately large.

There is generally a trade-off between the accuracy of the sample in representing population values and the costs associated with sample size. The larger the sample, the more confident we can be that it accurately reflects what exists in the population, but large samples can be extremely expensive and time consuming. A small sample is less expensive and time consuming, but it is not as accurate. Therefore, in situations requiring minimal error and maximum accuracy of prediction of population values, large samples will be required. In cases where more error can be tolerated, small samples will do. It is not unusual to use relatively small samples to generalize to millions of individuals. For example, a dramatic difference in size between the sample and the population to which the results are generalized is found in the Nielsen ratings. The Nielsen organization surveys the television viewing habits of about 5,000 households (13,000 persons) each week. From this survey, they make estimates concerning the viewing habits of around 250 million individuals watching television in the United States. This amounts to sampling about 1 in every 20,000 viewers. Even so, apparently the Nielsen estimates are sufficiently accurate to satisfy the competitive TV networks. This survey again illustrates that useful estimates of population values can be made using a small number of observations. The key factor in this case is the representativeness of the sample.

If an entire population were surveyed, then we could specify the makeup of that population—number of Democrats, Republicans, Independents, income, sex, and so on. However, to do this for the entire population of the United States would cost hundreds of millions of dollars. Once every ten years, the U.S. Census is taken. The cost of the last Census (2000) has been estimated at more than \$7 billion. For different purposes, far more limited than the U.S. Census, we could obtain a very good idea of some U.S. population characteristics from a relatively small representative sample of 2,000 to 3,000 people. That is, with a sample of this size, we could generalize our results to the entire population. In fact, some have argued that sample surveys can produce more accurate results regarding a given population than can a census. A census requires skilled interviewers in greater numbers than available. It also requires that *all* members of a population be reached. This does not always happen, and considerable unhappiness resulted in recent censuses because of what many viewed as an undercount of some segments of the population. Some cities were especially unhappy with the undercount because important funding decisions were tied to the census. A good sampling procedure might have been both less expensive and more accurate.

Other factors that help determine what is considered an adequate sample size are diversity of the population concerning the factors of interest and the number of factors. The greater the diversity among individuals and the greater the number of factors present, the larger the sample that is required to achieve representativeness.

Members of any population vary among themselves in many ways. In some instances, this variability may be small; in other instances, it may be quite large. To illustrate, say that we are interested in selecting a representative sample for height of individuals in a given population. If everyone in the given

population were identical in height—say, 6 feet tall—then a sample with one individual would be representative of the population. Even if our population varied slightly, say from 5 feet 11 inches to 6 feet 1 inch, a very small sample could be taken whose average would be representative of the population average. In fact, a sample with one individual, in this case, would not seriously misrepresent the population value. However, if the height of individuals in the population varied from 4 feet to 7 feet, our sample would have to be relatively large to obtain an average that was representative of the population average. In this instance, a small sample could seriously misrepresent the population average.

Sampling Error

Error can occur during the sampling process. **Sampling error** can include both systematic sampling error and random sampling error. **Systematic sampling error** is the fault of the investigation, but **random sampling error** is not. When errors are systematic, they bias the sample in one direction. Under these circumstances, the sample does not truly represent the population of interest. Systematic error occurs when the sample is not drawn properly, as in the poll conducted by *Literary Digest* magazine. It can also occur if names are dropped from the sample list because some individuals were difficult to locate or uncooperative. Individuals dropped from the sample could be different from those retained. Those remaining could quite possibly produce a biased sample. Political polls often have special problems that make prediction difficult.

Random sampling error, as contrasted to systematic sampling error, is often referred to as *chance error*. Purely by chance, samples drawn from the same population will rarely provide identical estimates of the population parameter of interest. These estimates will vary from sample to sample. For example, if you were to flip 100 unbiased coins, you would not be surprised if you obtained 55 heads on one trial, 49 on another, 52 on a third, and so on. Thus, some samples will, by chance, provide better estimates of the parameter than others. In any given sample, some attributes of interest may be overrepresented and some underrepresented. However, this type of error is random. Moreover, it is possible to describe this error statistically and take it into account when drawing inferences. Thus, in sampling voters prior to an election, we can make claims of the following sort: “There is a 95% probability that the proportion of voters who will cast their ballot for Candidate A will fall within the interval from 43% to 47%.” Before selecting a sample, researchers usually decide the amount of error they are willing to tolerate and the level of confidence they want to have. This margin of error is expressed in terms of a **confidence interval**—for example, 95% of the time the sample will correctly reflect the population values with a margin of error of plus or minus 4 percentage points.

Sampling error can affect inferences based on sampling in two important situations. In one situation, we may wish to generalize from the sample to a particular population. With a small sampling error, we can feel more confident that our sample is representative of the population. Thus, we can feel reasonably

comfortable about generalizing from the sample to the population. Survey research is most concerned about this kind of sampling error. The second situation in which sampling error plays a role is when we wish to determine whether two or more samples were drawn from the same or different populations. In this case, we are asking if two or more samples are sufficiently different to rule out factors due to chance. An example of this situation is when we ask the question “Did the group that received the experimental treatment really differ from the group that did not receive the treatment other than on the basis of chance?”

Evaluating Information From Samples

The information on sampling presented in this chapter not only guides the beginning researcher, it also guides all of us in being critical consumers of information based on sample data. In recent years, our exposure to sampling and information from samples has grown tremendously, and we need, more than ever, to be able to critically evaluate that information. Through the educational system and the media, we continue to be informed about new research findings from scientists around the world. We can be fairly confident of that information because scientists are trained in proper methodology and findings are generally reviewed by other scientists before the findings become public. Likewise, there are reputable organizations that gather information for the public. For example, the Gallup Organization specializes in conducting polls to assess attitudes. In the months leading to a presidential election, we are inundated on a daily basis with poll numbers that show the public’s views on a variety of political issues. Gallup polls are often used because the organization has a solid reputation for its sampling methodology and clearly states the margin of error in its polls. But what about other sources of information that have become more prevalent?

Polling and surveys are now routinely conducted in magazines, on radio talk shows, on Internet sites, and through 900 numbers on television programs. First, it is important to realize that none of these uses random sampling. All of the respondents who provide data self-select themselves. As a result, the sample of respondents often represents a small and narrow segment of the population. For example, the CNN Web site often has an ongoing poll regarding some issue in the news. When you view the results of the poll, there is an appropriate warning that reads, “This QuickVote is not scientific and reflects the opinions of only those Internet users who have chosen to participate. The results cannot be assumed to represent the opinions of Internet users in general, nor the public as a whole. The QuickVote sponsor is not responsible for content, functionality or the opinions expressed therein.” Unfortunately, many sources of polling do not provide such a warning to the consumer. We must realize that the results of a magazine survey reflect the views of only the individuals who read that type of magazine, have that magazine available to them to buy, actually buy that magazine, are interested enough in the survey to respond, have

the time to respond, and take the time to mail their answers. Exactly who are you left with in the sample, and what do the results mean? These are important questions to ask yourself as a consumer.

Case Analysis

On April 20, 1999, Eric Harris and Dylan Klebold entered Columbine High School and began shooting teachers and students. Thirteen individuals died, and the psychological community was again asked to explain such violent behavior. A psychologist might decide to interview Columbine students to obtain their perspectives on the factors that motivated the two young men to commit such a horrendous act. A group of ten students has already decided to meet and discuss the events of that day. The psychologist asks if it would be all right to attend the meeting and ask them some questions. The students agree, and the psychologist records their thoughts. Based on this information, the psychologist concludes that a primary reason for the violent behavior was the peer dynamics in the school that created groups of outcasts.

Critical Thinking Questions

1. Who appears to constitute the population of interest?
2. Which type of sampling procedure best describes that used by the psychologist?
3. What are the limitations of this sampling method, and in what specific ways could the sampling method have affected the findings?
4. What specific steps would you have taken to obtain a representative sample?

General Summary

When we conduct research, we are generally interested in drawing some conclusion about a population of individuals that have some common characteristic. However, populations are typically too large to allow observations on all individuals, and we resort to selecting a sample. In order to make inferences about the population, the sample must be representative. Thus, the manner in which the sample is drawn is critical. Probability sampling uses random sampling in which each element in the population (or a subgroup of the population with stratified random sampling) has an equal chance of being selected for the sample. This technique is considered to be the best means of obtaining a representative sample. When probability sampling is not possible, nonprobability sampling must be used. Convenience sampling involves using participants who are readily available (such as introductory psychology students). It is the easiest technique but the poorest from a methodological standpoint. Quota sampling is essentially convenience sampling in which there is an effort to better represent the population by sampling a certain percentage of participants from subgroups that correspond to the prevalence of those subgroups in the population.

By their very nature, samples do not perfectly match the population from which they are drawn. There is always some degree of sampling error, and the degree of error is inversely related to the size of the sample. Larger samples are more likely to accurately represent characteristics of the population, and smaller samples are less likely to accurately represent characteristics of the population. Therefore, researchers strive for samples that are large enough to reduce sampling error to an acceptable level. Even when samples are large enough, it is important to evaluate the specific method by which the sample was drawn. We are increasingly exposed to information obtained from self-selected samples that represent only a very narrow subgroup of individuals. Much of such information is meaningless because the subgroup is difficult to identify.

Detailed Summary

1. Sampling is the process whereby some elements (individuals) in the population are selected for a research study.
2. The population consists of all individuals with a particular characteristic that is of interest to the researchers. If data are obtained from all members of the population, then we have a census; if data are obtained from some members of the population, then we have a sample.
3. With probability sampling, a researcher can specify the probability of an element's (participant's) being included in the sample. With nonprobability sampling, there is no way of estimating the probability of an element's being included in a sample.
4. Although often more difficult and expensive, probability sampling is a methodologically more precise method to obtain a sample that is representative of the population.
5. With simple random sampling, each individual in the population has an equal chance of being selected for the sample. The four steps of simple random sampling are (1) defining the population, (2) constructing a list of all members, (3) drawing the sample, and (4) contacting the members of the sample.
6. Stratified random sampling is a form of probability sampling in which individuals are randomly selected from specified subgroups (strata) of the population. This method can be used to increase the representativeness of the sample and/or to allow comparisons to be made among individuals in the different strata.
7. Convenience sampling is quick and inexpensive because it involves selecting individuals who are readily available at the time of the study (such as introductory psychology students). The disadvantage is that convenience samples are generally less representative than random samples; therefore, results should be interpreted with caution.

8. Quota sampling involves the selection of a certain percentage of individuals from specified subgroups of the population when the population is large and lists of members are not available. Many polling organizations use this technique.
9. Appropriate sample size depends on various considerations, including population variability, statistical issues, economic factors, and availability of participants. In general, with larger samples you will have a smaller margin of error and you can detect smaller differences.
10. The larger the variability of scores in the population, the larger the sample must be in order to be representative.
11. Sampling error includes systematic error and random error. Systematic error occurs when the sample is not properly drawn (an error of the researcher). Random error is the degree to which the sample is not perfectly representative of the population. Even with the best sampling techniques, some degree of random error is expected.
12. Increasingly, we are exposed to information based on sample data. Understanding the principles of sampling, particularly the limitations of various methods, should make us more critical consumers of such information.

Key Terms

census

confidence interval

convenience sampling

element

exhaustive list

nonprobability sampling

population

probability sampling

quota sampling

random sampling

random sampling error

sample

sampling error

stratified random sample

systematic sampling error

Review Questions / Exercises

1. You hypothesize that students in fraternities/sororities on your campus are more extroverted (on an introversion/extroversion scale) than other students. To answer your research question, you need to obtain a sample of “Greek” students and a sample of “independent” students. For each of these two populations of students, how would you obtain simple random samples using the four steps described in this chapter?
2. For the hypothesis in question 1, describe how you would obtain samples using stratified random sampling, with stratification based on the sex of the student. Assume that 60% of the Greeks are women.
3. For the hypothesis in question 1, describe how you would obtain samples using convenience sampling.
4. For the hypothesis in question 1, describe how you would obtain samples using quota sampling based on the sex of the student. Again, assume that 60% of the Greeks are women.
5. For the hypothesis in question 1, describe a poor method of sampling that you believe would result in systematic sampling error and thus bias the research findings. Explain why you believe the sample would be biased.
6. Go to an Internet site that has an online poll. (If you are not familiar with such a site, simply conduct a search using the keywords “online poll” or “online survey” and select one of the sites.) Participate in the survey, and view the current results. Describe the nature of the survey and the current results. What method of sampling was used? What characteristics do you believe the members of the sample possess? Do you believe that these characteristics influence the poll results? Would the results be different if the general population of U.S. adults were polled?