Pearson New International Edition

Introduction to Behavioral Research Methods

Leary

Sixth Edition

Pearson New International Edition

Introduction to Behavioral Research Methods
Leary
Sixth Edition

PEARSON

somewhat from the characteristics of the general population. This difference, called **sampling error**, causes results obtained from a sample to differ from what would have been obtained had the entire population been studied. If you calculate the average grade point average of a representative sample of 200 students at your college or university, the mean for this sample will not perfectly match the average that you would obtain if you had used the grade point averages of *all* students in your school. If the sample is truly representative, however, the value obtained on the sample should be very close to what would be obtained if the entire population were studied.

Fortunately, when probability sampling techniques are used, researchers can estimate how much their results are affected by sampling error. The **error of estimation** (also called the **margin of error**) indicates the degree to which the data obtained from the sample are expected to deviate from the population as a whole. For example, you may have heard newscasters report the results of a political opinion poll and then add that the results "are accurate within 3 percentage points." What this means is that if 45% of the respondents in the sample endorsed Smith for president, we know that there is a 95% probability that the true percentage of people in the population who support Smith is between 42% and 48% (that is, 45% ± 3%). By allowing researchers to estimate the sampling error in their data, probability samples permit them to specify how confident they are that the results obtained on the sample accurately reflect the behavior of the population. Their confidence is expressed in terms of the error of estimation.

The smaller the error of estimation, the more closely the results from the sample estimate the behavior of the larger population. For example, if the limits on the error of estimation are only ±1%, the sample data are a better indicator of the population than if the limits on the error of estimation are ±10%. So, if the error of estimation in the opinion poll was 1%, we are rather confident that the true population value falls between 44% and 46% (that is, 45% ± 1%). But if the error of estimation is 10%, the true population has a 95% probability of being anywhere between 35% and 55% (that is, 45% ± 10%). Obviously, researchers prefer the error of estimation to be as small as possible.

The error of estimation is a function of three things: the sample size, the population size, and the variance of the data. First, the larger a probability sample, the more similar the sample tends to be to the population (that is, the smaller the sampling error) and the more accurately the sample data estimate the population's characteristics. You would estimate the average grade point average at your school more closely with a sample of 400 than with a sample of 50, for example, because larger sample sizes have a lower error of estimation.

The error of estimation also is affected by the size of the population from which the sample was drawn. Imagine we have two samples of 200 respondents. The first was drawn from a population of 400, the second from a population of 10 million. Which sample would you expect to mirror more closely the population's characteristics? I think you can guess that the error of estimation will be lower when the population contains 400 cases than when it contains 10 million cases.

The third factor that affects the error of estimation is the variance of the data. The greater the variability in the data, the more difficult it is to estimate the population values accurately. The larger the variance, the less representative the mean is of the set of scores as a whole. As a result, the larger the variance in the data, the larger the sample needs to be to draw accurate inferences about the population.

The error of estimation is meaningful only when we have a **probability sample**—a sample for which the researcher knows the mathematical probability that any individual in the population is included in the sample. Only with a probability sample do we know that the statistics that we calculate from the sample data reflect the true values in the parent population, at least within the margin defined by the error of estimation. If we do not have a probability sample, the characteristics of the sample may not reflect those of the population, so we cannot trust that the sample statistics tell us anything at all about the population. In this case, the error of estimation is irrelevant because the data cannot be used to draw inferences about the population anyway.

Thus, when researchers want to draw inferences about a population from a sample, they must select a probability sample. Probability samples may be obtained in several ways, but four basic methods involve simple random sampling, systematic sampling, stratified random sampling, and cluster sampling.

Simple Random Sampling

When a sample is chosen in such a way that every possible sample of the desired size has the same chance of being selected from the population, the sample is a **simple random sample**. For example, suppose we want to select a sample of 200 participants from a school district that has 5,000 students. If we wanted a simple random sample, we would select our sample in such a way that every possible combination of 200 students has the same probability of being chosen.

To obtain a simple random sample, the researcher must have a **sampling frame**—a list of the population from which the sample will be drawn. Then participants are chosen randomly from this list. If the population is small, one approach is to write the name of each case in the population on a slip of paper, shuffle the slips of paper, then pull slips out until a sample of the desired size is obtained. For example, we could type each of the 5,000 students' names on cards, shuffle the cards, then randomly pick 200. However, with larger populations, pulling names "out of a hat" becomes unwieldy.

The primary way that researchers select a random sample is to number each person in the sampling frame from 1 to N, where N is the size of the population. Then they pick a sample of the desired size by selecting numbers from 1 to N by some random process. Traditionally, researchers have used a **table of random numbers**, which contains long rows of numbers that have been generated in a random order. (Tables of random numbers can be found in many statistics books and on the Web.) Today, researchers more commonly use computer programs to generate lists of random numbers, and you can find Web sites that allow you to generate lists of random numbers from 1 to whatever sized population you might have. Whether generated from a table or by a computer, the idea is the same. Once we have numbered our sampling frame from 1 to N and generated as many random numbers as needed for the desired sample size, the individuals in our sampling frame who have the randomly generated numbers are selected for the sample.

In Depth

Random Telephone Surveys: The Problem of Cell Phones

Not too many years ago, almost all American households had a single telephone line. As a result, phone numbers provided a convenient sampling frame from which researchers could choose a random sample of households for surveys. Armed with a population of phone numbers, researchers could easily draw a random sample. Although researchers once used phone books to select their samples, for the past few decades they have relied upon random digit dialing. **Random digit dialing** is a method for selecting a random sample for telephone surveys by generating telephone numbers at random. Random digit dialing is better than choosing numbers from a phone book because it will generate unlisted numbers as well as listed ones.

However, the spread of cell phones has created a number of problems for researchers who rely on random digit dialing to obtain random samples. First, the Telephone Consumer Protection Act prohibits using an automatic dialer to call cell phone numbers. Researchers could dial them manually, but then the advantages of using automated dialing are lost. Second, because many households have both a landline and one or more cell phones, households differ in the likelihood that they will be contacted for the study. (Households with more phone numbers are more likely to be sampled.) As we saw earlier, a probability sample requires that researchers estimate the probability that a particular case will be included in the sample, but this is not possible if households differ in the number of phones they have. Third, researchers often want to confine their probability sample to a particular geographical region—a particular city or state, for example. But because people can keep their cell phone number when they move, the area code for a cell phone does not reflect the person's location as it does with landline phone numbers. Finally, people may be virtually anywhere when they answer their cell phone. Researchers worry that the quality of the data they collect as people are driving, standing in line, shopping, sitting in the bathroom, visiting, and multitasking in other ways is not as good as when people are in the privacy of their own homes (Keeter, Kennedy, Clark, Tompson, & Mokrzycki, 2007; Link, Battaglia, Frankel, Osborn, & Mokdad, 2007).

On top of these methodological issues, evidence suggests that people who use only a cell phone differ on average from those who have only a landline phone or both landlines and cell phones. This fact was discovered during the 2004 presidential election when phone surveys underestimated the public's support for John Kerry in his campaign against George W. Bush. The problem arose because people who had only a cell phone (but no landline phone) were more likely to support Kerry than those who had landline phones. Not only do they differ in their demographic characteristics (for example, they are younger and more likely to be unmarried), but they hold different political attitudes, watch different TV shows, and are more likely to use computers to get the news. Not surprisingly, then, the results of cell phone surveys often differ from the results of landline phone surveys. And to make matters worse, among people who have both cell phones and landlines, those who are easier to reach on their cell phone differ from those who are easier to reach on their landline phone at home (Link et al., 2007). Fortunately, because the number of people who have a cell phone but no landline home phone remains small, using random digit dialing to contact people with landline phones may not influence the results of telephone surveys much for now (Keeter et al., 2007). But as the number of cell phones grows and home-based landline phones continue to disappear, researchers will need to find new ways to grapple with this problem.

Systematic Sampling

One major drawback of simple random sampling is that we must know how many individuals are in the population and have a sampling frame that lists all of them before we begin. Imagine that we wish to study people who use hospital emergency rooms for psychological rather than medical problems. We cannot use simple random sampling because at the time that we start the study, we have no idea how many people might come through the emergency room during the course of the study and don't have a sampling frame. In such a situation, we might choose to use **systematic sampling**. Systematic sampling involves taking every so many individuals for the sample. For example, we could decide that we would interview every 8th person who came to the ER for care until we obtained a sample of whatever size we desired. When the study is over, we will know how many people came through the emergency room and how many we

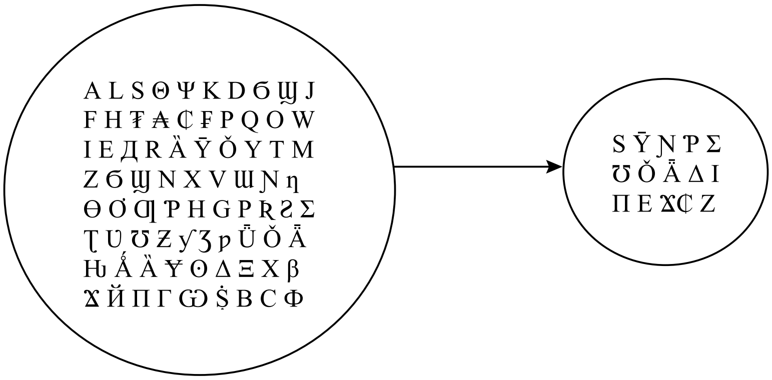

FIGURE 1 Simple Random Sampling. In this figure, the population is represented by the large circle, the sample is represented by the small circle, and the letters are individual people. In simple random sampling, cases are sampled at random directly from the population in such a way that every possible sample of the desired size has an equal probability of being chosen.

selected and, thus, we would also know the probability that any person who came to the ER during the study would be in our sample.

You may be wondering why this is not a simple random sample. The answer is that, with a simple random sample, every possible sample of the desired size has the same chance of being selected from the population. In systematic sampling this is not the case. After we select a particular participant, the next several people have no chance at all of being in the sample. For example, if we are selecting every 8th person for the study, the 9th through the 15th persons to walk into the ER have no chance of being chosen, and our sample could not possibly include, for example, both the 8th and the 9th person. In a simple random sample, all possible samples have an equal chance of being used, so this combination would be possible.

Stratified Random Sampling

Stratified random sampling is a variation of simple random sampling. Rather than selecting cases directly from the population, we first divide the population into two or more subgroups or strata. A **stratum** is a subset of the population that shares a particular characteristic. For example, we might divide the population into men and women, into different racial groups, or into six age ranges (20–29, 30–39, 40–49, 50–59, 60–69, over 69). Then cases are randomly sampled from each of the strata.

Stratification ensures that researchers have adequate numbers of participants from each stratum so that they can examine differences in responses among the various strata. For example, the researcher might want to compare younger respondents (20–29 years old) with older respondents (60–69 years old). By first stratifying the sample, the researcher ensures that there will be an ample number of both young and old respondents in the sample.

In many cases, researchers use a **proportionate sampling method** in which cases are sampled from each stratum in proportion to their prevalence in the population. For example, if the registered voters in a city are 55% Democrats and 45% Republicans, a researcher studying political attitudes may wish to sample proportionally from those two strata to be sure that the sample is also composed of 55% Democrats and 45% Republicans. When this is done, stratified random sampling can increase the probability that the sample we select will be representative of the population.

Cluster Sampling

Although they provide us with very accurate pictures of the population, simple and stratified random sampling have a major drawback: They require that we have a sampling frame of all cases in the population before we begin. Obtaining a list of small, easily identified populations is no problem. You would find it relatively easy to obtain a list of all students in your

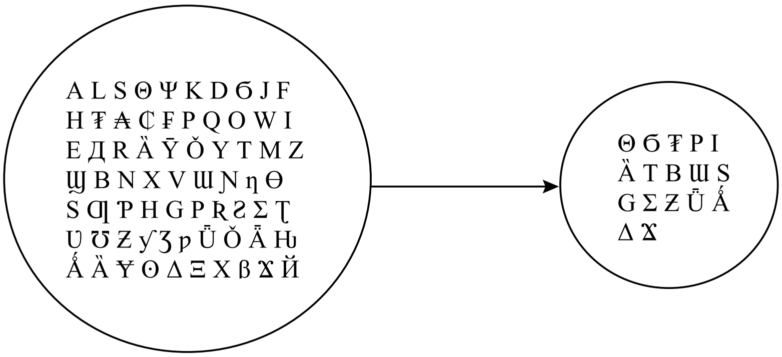

FIGURE 2 Systematic Sampling. In this figure, the population is represented by the large circle, the sample is represented by the small circle, and the letters are individual people. In systematic sampling, every *n*th person is selected from a list. In this example, every 4th person has been chosen.

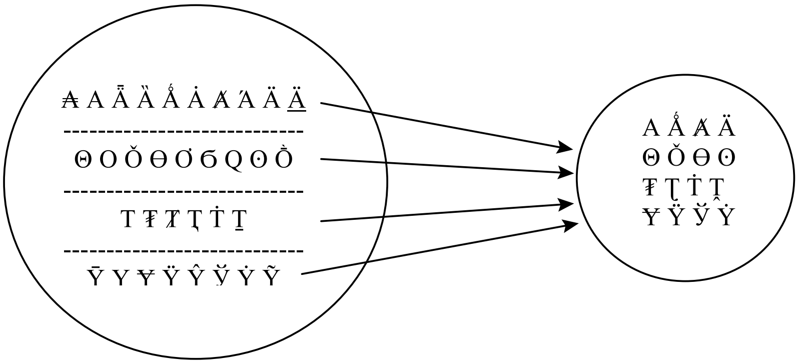

FIGURE 3 Stratified Random Sampling. In this figure, the population is represented by the large circle, the sample is represented by the small circle, and the letters are individual people. In stratified random sampling, the population is first divided into strata composed of individuals who share a particular characteristic. In this example, the population is divided into four strata. Then cases are randomly selected from each of the strata.

college or all members of the Association for Psychological Science, for example. Unfortunately, not all populations are easily identified. Could we, for example, obtain a list of every person in the United States or, for that matter, in New York City or Miami? Could we get a sampling frame of all Hispanic 3-year-olds, all people who are deaf who know sign language, or all single-parent families in Canada headed by the father? In cases such as these, random sampling is not possible because without a list we cannot locate potential participants or specify the probability that a particular case will be included in the sample.

In such instances, **cluster sampling** is often used. To obtain a cluster sample, the researcher first samples not participants but rather groupings or *clusters* of participants. These clusters are often based on naturally occurring groupings, such as geographical areas or particular institutions. For example, if we wanted a sample of elementary school children in West Virginia, we might first randomly sample from the 55 county school systems in West Virginia. Perhaps we would pick 15 counties at random. Then, after selecting this small random sample of counties, we could get lists of students for those counties and obtain random samples of students from the selected counties.

Often cluster sampling involves a **multistage cluster sampling** process in which we begin by sampling large clusters, then we sample smaller clusters from within the large clusters, then we sample even smaller clusters, and finally we obtain our sample of participants. For example, we could randomly pick counties and then randomly choose several particular schools from the selected counties. We could then randomly select particular classrooms from the schools we selected, and finally randomly sample students from each classroom.

Cluster sampling has two advantages. First, a sampling frame of the population is not needed to begin sampling—only a list of the clusters. In this example, all we would need to start is a list of counties in West Virginia, a list that would be far easier to obtain than a census of all children enrolled in West Virginia schools. Then, after sampling the clusters, we can get lists of students within each cluster (that is, county) that was selected, which is much easier than getting a list of the entire population of students in West Virginia. The second advantage is that, if each cluster represents a grouping of participants that are close together geographically (such as students in a certain county or school), less time and effort are required to contact the participants. Focusing on only 15 counties would require considerably less time, effort, and expense than sampling students from all 55 counties in the state.

In Depth

To Sample or Not to Sample: The Census Debate

Since the first U.S. census in 1790, the Bureau of the Census has struggled to find ways to account for every person in the country. For a variety of reasons, many citizens are miscounted by census-takers. The population of the United States is not only large, but it is also moving, changing, and partially hidden, and any effort to count the entire population will both overcount and undercount certain groups. In the 2000 census, for example, an estimated 6.4 million people were not counted, and approximately 3.1 million people appear to have been counted twice. The challenge that faces the Census Bureau is to design and administer the census in a way that provides the most accurate data. To do so, the Census Bureau has proposed to rely on sampling procedures rather than to try to track down each and every person.

The big problem that compromises the validity of the census is that a high percentage of people either do not receive the census questionnaire or, if they receive it, do not complete and return it as required by law. So, how can we track these nonresponders down? Knowing that it will be impossible to visit every one of the millions of households that did not respond to the mailed questionnaire or follow-up call, the bureau proposed that census-takers visit a ***representative sample*** of the addresses that do not respond. The rationale is that, by focusing their time and effort on this representative sample rather than trying to contact every household that is unaccounted for (which previous censuses showed is fruitless), they could greatly increase their chances of obtaining the missing information from these otherwise uncounted individuals. Then, using the data from the representative sample of nonresponding households, researchers could estimate the size and demographic characteristics of other missing households. Once they know the racial, ethnic, gender, and age composition of this representative sample of people who did not return the census form, statistical models can be used to estimate the characteristics of the entire population that did not respond.

Statisticians overwhelmingly agree that sampling will dramatically improve the accuracy of the census. A representative sample of nonresponding individuals provides far more accurate data than an incomplete set of households that is biased in unknown ways. However, despite its statistical merit, the plan met stiff opposition in Congress, and the Supreme Court ruled that sampling techniques could not be used to reapportion seats in the House of Representatives. Many people have trouble believing that contacting a probability sample of nonresponding households provides far more accurate data than trying (and failing) to locate them all, although you should now be able to see that this is the case. In addition, many politicians worry that the sample would be somehow biased (resulting perhaps in loss of federal money to their districts), would underestimate members of certain groups, or would undermine public trust in the census. Such concerns reflect misunderstandings about probability sampling.

Despite the fact that sampling promised to both improve the accuracy of the census and lower its cost, Congress denied the Census Bureau's request to use sampling in the 2000 and 2010 census. However, although the bureau was forced to attempt a full-scale enumeration of every individual in the country (a challenge that was doomed to failure from the outset), it was allowed to study sampling procedures to document their usefulness. Unfortunately, politics have prevailed over reason and science, and opponents have blocked the use of sampling procedures that would undoubtedly provide a better estimate of the population's characteristics.

The Problem of Nonresponse

The **nonresponse problem** is the failure to obtain responses from individuals that researchers select for a sample. In practice, researchers are rarely able to obtain perfectly representative samples because some people who are initially selected for the sample either cannot be contacted or refuse to participate. For example, when households or addresses are used as the basis of sampling, interviewers may repeatedly find that no one is at home when they visit the address. Or, in the case of mailed surveys, the person selected for the sample may have moved and left no forwarding address. If the people who can easily be located differ from those who cannot be found, the people who can be found may not be representative of the population as a whole and the results of the study may be biased in unknown ways.

Even when people who are selected for the sample are contacted, a high proportion of them do not