

Sample Surveys in the Real World

4

CASE STUDY An opinion poll talks to 1000 people chosen at random, announces its results, and announces a margin of error. Should we be happy? Perhaps not. Many polls don't tell the whole truth about their samples. The Pew Research Center for the People and the Press did tell the whole truth about its sampling methods. Here it is.

Most polls are taken by telephone, dialing numbers at random to get a random sample of households. After eliminating fax and business numbers, Pew had to call 2879 residential numbers to get its sample of 1000 people. Here's the breakdown:

Never answered phone	938
Answered but refused	678
Not eligible: no person aged 18 or older, or language barrier	221
Incomplete interview	42
Complete interview	1000
Total called	2879



Tetra Images/SuperStock

Out of 2879 working residential phone numbers, 33% never answered. Of those who answered, 24% refused to talk. The overall rate of nonresponse (people who never answered, refused, or would not complete the interview) was 1658 out of 2879, or 58%. Pew called every number five times over a five-day period, at different times of day and on different days of the week. Many polls call only once, and it is usual to find that more than half of those who answer refuse to talk. Although Pew did obtain the desired sample of 1000 people, can we trust the results of this poll to make conclusions about people with residential phones? Can the results of the poll be extended to people who only have cell phones? By the end of this chapter, you will learn how to answer this question.

How sample surveys go wrong

Random sampling eliminates bias in choosing a sample and allows control of variability. So once we see the magic words “randomly selected” and “margin of error,” do we know we have trustworthy information before us? It certainly beats voluntary response, but not always by as much as we might hope. Sampling in the real world is more complex and less reliable than choosing an SRS from a list of names in a textbook exercise. Confidence statements do not reflect all the sources of error that are present in practical sampling.

Errors in sampling

Sampling errors are errors caused by the act of taking a sample. They cause sample results to be different from the results of a census.

Random sampling error is the deviation between the sample statistic and the population parameter caused by chance in selecting a random sample. The margin of error in a confidence statement includes *only* random sampling error.

Nonsampling errors are errors not related to the act of selecting a sample from the population. They can be present even in a census.

Most sample surveys are afflicted by errors other than random sampling errors. These errors can introduce bias that makes a confidence statement meaningless. Good sampling technique includes the art of reducing all sources of error. Part of this art is the science of statistics, with its random samples and confidence statements. In practice, however, good statistics isn't all there is to good sampling. Let's look at sources of errors in sample surveys and at how samplers combat them.

Sampling errors

Random sampling error is one kind of sampling error. The margin of error tells us how serious random sampling error is, and we can control it by choosing the size of our random sample. Another source of sampling error is the use of *bad sampling methods*, such as voluntary response. We can avoid bad methods. Other sampling errors are not so easy to handle. Sampling begins with a list of individuals from which we will draw our sample. This list is called the **sampling frame**. Ideally, the sampling frame should list every individual in the population. Because a list of the entire population is rarely available, most samples suffer from some degree of *undercoverage*.

Undercoverage

Undercoverage occurs when some groups in the population are left out of the process of choosing the sample.

If the sampling frame leaves out certain classes of people, even random samples from that frame will be biased. Using telephone directories as the frame for a telephone survey, for example, would miss everyone with an unlisted residential telephone number, everyone who cannot afford a residential phone, and everyone who has only a cell phone. Even if we consider only those with residential telephones, more than half the households in many large cities have unlisted numbers, so massive undercoverage and bias against urban areas would result. In fact, telephone surveys use random digit dialing equipment, which dials telephone numbers in selected regions at random. In effect, the sampling frame contains all residential telephone numbers.



Opting out of modern society was tiring at times, but Ted felt that increasing the undercoverage of opinion polls made it all worthwhile.

EXAMPLE 1 Did we miss anyone?

Most opinion polls can't afford even to attempt full coverage of the population of all adult residents of the United States. The interviews are done by telephone, thus missing the 2% of households without phones. Only *households* are contacted, so students in dormitories, prison inmates, and most members of the armed forces are left out. So are the homeless and people staying in shelters. Many polls interview only in English, which leaves some immigrant households out of their samples.

The kinds of undercoverage found in most sample surveys are most likely to leave out people who are young or poor or who move often. Nonetheless, random digit dialing comes close to producing a random sample of households with phones. Sampling errors in careful sample surveys are usually quite small. The real problems start when someone picks up (or doesn't pick up) the phone. Now nonsampling errors take over.

Nonsampling errors

Nonsampling errors are those that can plague even a census. They include **processing errors**—mistakes in mechanical tasks such as doing arithmetic or entering responses into a computer. Fortunately, the wide availability and adoption of computers has made processing errors less common than in the past.

EXAMPLE 2 Computer-assisted interviewing

The days of the interviewer with a clipboard are past. Contemporary interviewers carry a laptop computer for face-to-face interviews or watch a computer screen as they carry out a telephone interview. Computer software manages the interview. The interviewer reads questions from the computer screen and uses the keyboard to enter the responses. The computer skips irrelevant items—once a respondent says that she has no children, further questions about her children never appear. The computer can check that answers to related questions are consistent with each other. It can even present questions in random order to avoid any bias due to always asking questions in the same order.

Computer software also manages the record keeping. It keeps records of who has responded and prepares a file of data from the responses. The tedious process of transferring responses from paper to computer, once a source of processing errors, has disappeared. The computer even schedules the calls in telephone surveys, taking account of the respondent's time zone and honoring appointments made by people who were willing to respond but did not have time when first called.

Another type of nonsampling error is **response error**, which occurs when a subject gives an incorrect response. A subject may lie about her age or income or about whether she has used illegal drugs. She may remember incorrectly when asked how many packs of cigarettes she smoked last week. A subject who does not understand a question may guess at an answer rather than appear ignorant. Questions that ask subjects about their behavior during a fixed time period are notoriously prone to response errors due to faulty memory. For example, the National Health Survey asks people how many times they have visited a doctor in the past year. Checking their responses against health records found that they failed to remember 60% of their visits to a doctor. A survey that asks about sensitive issues can also expect response errors, as the next example illustrates.

EXAMPLE 3 The effect of race

In 1989, New York City elected its first black mayor and the state of Virginia elected its first black governor. In both cases, samples of voters interviewed as they left their polling places predicted larger margins of victory than the official vote counts. The polling organizations were certain that some voters lied when interviewed because they felt uncomfortable admitting that they had voted against the black candidate. This phenomenon is known as “social desirability bias” and “the Bradley effect,” after Tom Bradley, the former black mayor of Los Angeles who lost the 1982 California gubernatorial election despite leading in final-day preelection polls.

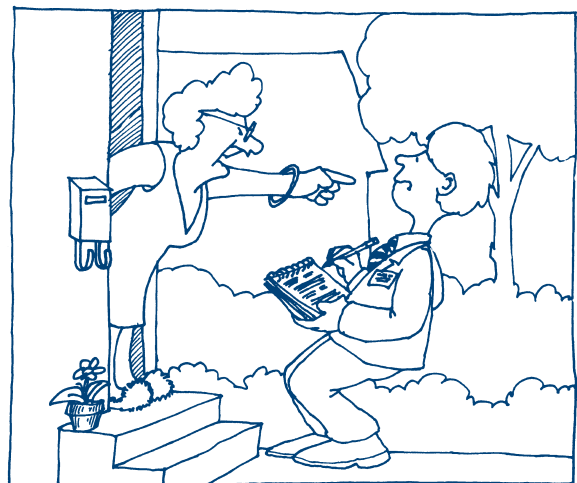
This effect attracted media attention during the 2008 presidential election. A few weeks before the election, polls predicted a victory, possibly a big one, for Barack Obama. Even so, Democrats worried that these polls might be overly optimistic because of the Bradley effect. In this case their fears were unfounded, but some political scientists claimed to detect the Bradley effect in polls predicting outcomes in primary races between Barack Obama and Hilary Clinton (for example, in the New Hampshire primary, polls predicted an 8 percentage point Obama victory, but Clinton won by 3 percentage points).

Technology and attention to detail can minimize processing errors. Skilled interviewers greatly reduce response errors, especially in face-to-face interviews. There is no simple cure, however, for the most serious kind of nonsampling error, *nonresponse*.

Nonresponse

Nonresponse is the failure to obtain data from an individual selected for a sample. Most nonresponse happens because some subjects can't be contacted or because some subjects who are contacted refuse to cooperate.

Nonresponse is the most serious problem facing sample surveys. People are increasingly reluctant to answer questions, particularly over the phone. The rise of telemarketing, voicemail, and caller ID drives down response to telephone surveys. Gated communities and buildings



"You can call, you can send email, you can stand at the door all day. The answer is still NO!"

guarded by doormen prevent face-to-face interviews. Nonresponse can bias sample survey results because different groups have different rates of nonresponse. Refusals are higher in large cities and among the elderly, for example. Bias due to nonresponse can easily overwhelm the random sampling error described by a survey's margin of error.

EXAMPLE 4 How bad is nonresponse?

The U.S. Census Bureau's American Community Survey (ACS) is a monthly survey of almost 300,000 housing units and replaced the U.S. Census Bureau's "long form" that, in the past, was sent to some households in the every-10-years national census. Participation in the ACS is mandatory, and the U.S. Census Bureau follows up by telephone and then in person if a household fails to return the mail questionnaire.

The ACS has the lowest nonresponse rate of any poll we know: in 2013, only about 1.3% of the households in the sample refused to respond; the overall nonresponse rate, including "never at home" and other causes, was 10.1%. This is a stark contrast in nonresponse from previous years—in 2012, the total nonresponse rate was only 2.7%. In October 2013, there was a government shutdown. During that time, "the ACS did not have a second mailing, a telephone followup, or a person followup operation for the October 2013 housing unit panel. . . . This caused a drop in the annual housing unit response rate of about 7 percentage points." If October 2013 is excluded from the ACS when considering nonresponse, the total nonresponse rate for 2013 was only 2.9%, similar to previous years.

Another survey that has a remarkable response rate is the University of Chicago's General Social Survey (GSS), the nation's most important social survey. The GSS (Example 7 in Chapter 1) contacts its sample in person, and it is run by a university. Despite these advantages, its recent surveys have a 30% rate of nonresponse.

What about polls done by the media and by market research and opinion-polling firms? We often don't know their rates of nonresponse because they won't say. That itself is a bad sign. The Pew poll we looked at in the Case Study suggests how bad things are. Pew got 1221 responses (of whom 1000 were in the population they targeted) and 1658 who were never at home, refused, or would not finish the interview. That's a nonresponse rate of 1658 out of 2879, or 58%. The Pew researchers were more thorough than many pollsters. Insiders say that nonresponse often reaches 75% or 80% of an opinion poll's original sample.

The situation investigated in the Case Study was more promising than a recent Pew survey. In a December 2013 survey on social media use, the Pew Research Center provided a full disposition of the sampled phone numbers. Here are the details. Initially, 40,985 landline and

27,000 cell phones were dialed, with 11,260 of the landlines and 15,758 of the cell numbers being working numbers. Of these 27,018 working numbers, they were able to contact 17,335, or about 64%, as a large portion of the calls went to voicemail. Among those contacted, about 15% cooperated. Of those cooperating, some numbers were ineligible due to language barriers or contacting a child's cell phone, and some calls were eventually broken off without being completed. To sum it up, the 27,018 working numbers dialed resulted in a final sample of 1801, giving a response rate of about 7%, an estimate of the fraction of all *eligible* respondents in the sample who were ultimately interviewed.

Sample surveyors know some tricks to reduce nonresponse. Carefully trained interviewers can keep people on the line if they answer at all. Calling back over longer time periods helps. So do letters sent in advance. Letters and many callbacks slow down the survey, so opinion polls that want fast answers to satisfy the media don't use them. Even the most careful surveys find that nonresponse is a problem that no amount of expertise can fully overcome. That makes this reminder even more important:

What the margin of error doesn't say

The announced margin of error for a sample survey covers only random sampling error. Undercoverage, nonresponse, and other practical difficulties can cause large bias that is not covered by the margin of error.

Careful sample surveys warn us about the other kinds of error. The Pew Research Center, for example, says, "In addition to sampling error, one should bear in mind that question wording and practical difficulties in conducting surveys can introduce error or bias into the findings of opinion polls." How true it is.

Does nonresponse make many sample surveys useless? Maybe not. We began this chapter with an account of a "standard" telephone survey done by the Pew Research Center. The Pew researchers also carried out a "rigorous" survey, with letters sent in advance, unlimited calls over eight weeks, letters by priority mail to people who refused, and so on. All this drove the rate of nonresponse down to 30%, compared with 58% for the standard survey. Pew then compared the answers to the same questions from the two surveys. The two samples were quite similar in age, sex, and race, though the rigorous sample was a bit more prosperous. The two samples also held



He started it! A study of deaths in bar fights showed that in 90% of the cases, the person who died started the fight. You shouldn't believe this. If you killed someone in a fight, what would you say when the police ask you who started the fight? After all, dead men tell no tales. Now that's nonresponse.

similar opinions on all issues except one: race. People who at first refused to respond were less sympathetic toward the plights of blacks and other minorities than those who were willing to respond when contacted the first time. Overall, it appears that standard polls give reasonably accurate results. But, as in Example 3, race is again an exception.

Wording questions

A final influence on the results of a sample survey is the exact wording of questions. It is surprisingly difficult to word questions so that they are completely clear. A survey that asks “Do you enjoy watching football?” will generate different answers based on the respondent’s understanding of “football” (American football or soccer).

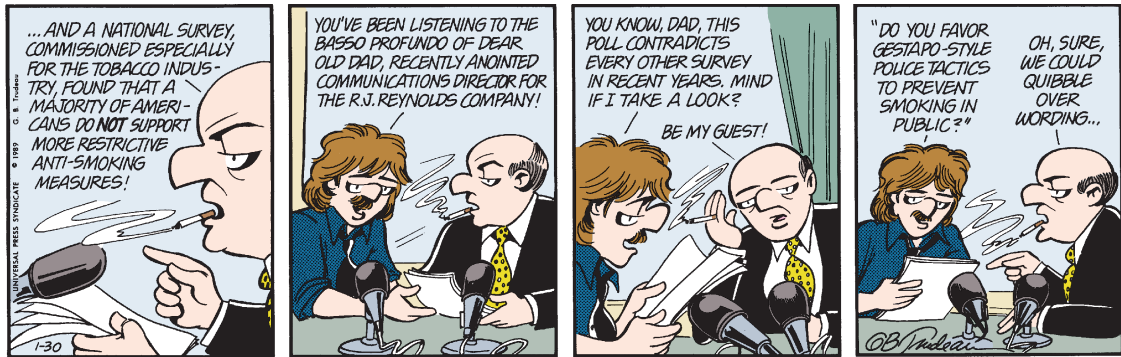
EXAMPLE 5 Words make a big difference

In May 2013, the Pew Research Center and the *Washington Post*/ABC News conducted polls asking whether people felt the U.S. Department of Justice had the right to subpoena Associated Press reporters’ phone records. Each survey phrased the question differently, and each survey found different results.

When the Pew Research Center asked, “Do you approve or disapprove of the Justice Department’s decision to subpoena the phone records of AP journalists as part of an investigation into the disclosure of classified information?” 36% of respondents approved. The *Washington Post*/ABC News survey said, “The AP reported classified information about U.S. anti-terrorism efforts and prosecutors have obtained AP’s phone records through a court order. Do you think this action by federal prosecutors is or is not justified?” Fifty-two (52%) of respondents to this survey said that the action of federal prosecutors was justified.

Before asking about approval or disapproval of the Justice Department’s actions, Pew asked how closely respondents had followed the issues that led up to the subpoena, and 64% reported not following the news story too closely. Additionally, the difference in wording—“decision to subpoena” in the Pew Research survey versus “obtained . . . through a court order” used by the *Washington Post*/ABC News survey—could have led to legitimizing the actions of the Department of Justice and thus a higher “approval” for the Justice Department’s actions. Or, was the higher justification based on the inclusion of the mention of U.S. antiterrorism efforts? Or, perhaps the difference is due to Pew using “the Justice Department” and the *Washington Post*/ABC News using “federal prosecutors.” We cannot begin to determine which part of the wording affected the responses. As the Pew Research Center said in its

Doonesbury



BY GARRY TRUDEAU

DOONESBURY © 1989 G. B. Trudeau. Reprinted with permission of UNIVERSAL UCLICK. All rights reserved.

article comparing the results of these two surveys, “each polling organization made good-faith efforts to describe the facts of the situation as accurately as possible, but the word choices and context make it impossible to identify one particular phrase or concept that tipped the public’s thinking.”

The wording of questions always influences the answers. If the questions are slanted to favor one response over others, we have another source of nonsampling error. A favorite trick is to ask if the subject favors some policy as a means to a desirable end: “Do you favor banning private ownership of handguns in order to reduce the rate of violent crime?” and “Do you favor imposing the death penalty in order to reduce the rate of violent crime?” are loaded questions that draw positive responses from people who are worried about crime. Here is another example of the influence of question wording.

EXAMPLE 6 Paying taxes

In April 2015, a Gallup Poll asked two questions about the amount one pays in federal income taxes. Here are the two questions:

Do you consider the amount of federal income tax you have to pay as too high, about right, or too low?

Do you regard the income tax which you will have to pay this year as fair?

The first question had 42% of respondents say “about right” (51% said “too high”), while the second question resulted in 56% of respondents saying that the taxes they paid were “fair.” There appears to be a definite difference in opinions about taxes when people are asked about the amount they have to pay or whether it is fair.

**NOW IT'S
YOUR TURN**

4.1 Should we recycle? Is the following question slanted toward a desired response? If so, how?

In view of escalating environmental degradation and incipient resource depletion, would you favor economic incentives for recycling of resource-intensive consumer goods?

How to live with nonsampling errors

Nonsampling errors, especially nonresponse, are always with us. What should a careful sample survey do about this? First, **substitute other households** for the nonresponders. Because nonresponse is higher in cities, replacing nonresponders with other households in the same neighborhood may reduce bias. Once the data are in, all professional surveys use statistical methods to **weight the responses** in an attempt to correct sources of bias. If many urban households did not respond, the survey gives more weight to those that did respond. If too many women are in the sample, the survey gives more weight to the men. Here, for example, is part of a statement in the *New York Times* describing one of its sample surveys:

The results have been weighted to take account of household size and number of telephone lines into the residence and to adjust for variations in the sample relating to geographic region, sex, race, age and education.

The goal is to get results “as if” the sample matched the population in age, gender, place of residence, and other variables.

The practice of weighting creates job opportunities for statisticians. It also means that the results announced by a sample survey are rarely as simple as they seem to be. Gallup announces that it interviewed 1523 adults and found that 57% of them bought a lottery ticket in the last 12 months. It would seem that because 57% of 1523 is 868, Gallup found that 868 people in its sample had played the lottery. Not so. Gallup no doubt used some quite fancy statistics to weight the actual responses: 57% is Gallup's best estimate of what it would have found in the absence of nonresponse. Weighting does help correct bias. It usually also increases variability. The announced margin of error must take this into account, creating more work for statisticians.

Sample design in the real world

The basic idea of sampling is straightforward: take an SRS from the population and use a statistic from your sample to estimate a parameter of the population. We now know that the sample statistic is altered behind the scenes to partly correct for nonresponse. The statisticians also have their hands on our beloved SRS. In the real world, most sample surveys use more complex designs.

EXAMPLE 7 The Current Population Survey

The population that the Current Population Survey (CPS) is interested in consists of all households in the United States (including Alaska and Hawaii). The sample is **chosen in stages**. The Census Bureau divides the nation into 2007 geographic areas called primary sampling units (PSUs). These are generally groups of neighboring counties. At the first stage, 754 PSUs are chosen. This isn't an SRS. If all PSUs had the same chance to be chosen, the sample might miss Chicago and Los Angeles. So 428 highly populated PSUs are automatically in the sample. The other 1579 are divided into 326 groups, called **strata**, by combining PSUs that are similar in various ways. One PSU is chosen at random to represent each stratum.

Each of the 754 PSUs in the first-stage sample is divided into census blocks (smaller geographic areas). The blocks are also grouped into strata, based on such things as housing types and minority population. The households in each block are arranged in order of their location and divided into groups, called **clusters**, of about four households each. The final sample consists of samples of clusters (not of individual households) from each stratum of blocks. Interviewers go to all households in the chosen clusters. The samples of clusters within each stratum of blocks are also not SRSs. To be sure that the clusters spread out geographically, the sample starts at a random cluster and then takes, for example, every 10th cluster in the list.

The design of the CPS illustrates several ideas that are common in real-world samples that use face-to-face interviews. Taking the sample *in several stages* with *clusters* at the final stage saves travel time for interviewers by grouping the sample households first in PSUs and then in clusters. Note that clustering is not an aspect of all sampling strategies, but it can be quite helpful in situations like the CPS.

The most important refinement mentioned in Example 7 is *stratified sampling*.

Stratified sample

To choose a **stratified random sample**:

Step 1. Divide the sampling frame into distinct groups of individuals, called **strata**. Choose the strata according to any special interest you have in certain groups within the population or because the individuals in each stratum resemble each other.

Step 2. Take a separate SRS in each stratum and combine these to make up the complete sample.

We must, of course, choose the strata using facts about the population that are known before we take the sample. You might group a university's students into undergraduate and graduate students or into those who live on campus and those who commute. Stratified samples have some advantages over an SRS. First, by taking a separate SRS in each stratum, we can set sample sizes to allow separate conclusions about each stratum. Second, a stratified sample usually has a smaller margin of error than an SRS of the same size. The reason is that the individuals in each stratum are more alike than the population as a whole, so working stratum by stratum eliminates some variability in the sample.

It may surprise you that stratified samples can violate one of the most appealing properties of the SRS—stratified samples need not give all individuals in the population the same chance to be chosen. Some strata may be deliberately overrepresented in the sample.

EXAMPLE 8 Stratifying a sample of students

A large university has 30,000 students, of whom 3000 are graduate students. An SRS of 500 students gives every student the same chance to be in the sample. That chance is

$$\frac{500}{30,000} = \frac{1}{60}$$

We expect an SRS of 500 to contain only about 50 grad students—because grad students make up 10% of the population, we expect them to make up about 10% of an SRS. A sample of size 50 isn't large enough to estimate grad student opinion with reasonable accuracy. We might prefer a stratified random sample of 200 grad students and 300 undergraduates.

You know how to select such a stratified sample. Label the graduate students 0001 to 3000 and use Table A to select an SRS of 200. Then label the undergraduates 00001 to 27000 and use Table A a second time to select an SRS of 300 of them. These two SRSs together form the stratified sample.

In the stratified sample, each grad student has chance

$$\frac{200}{3000} = \frac{1}{15}$$

to be chosen. Each of the undergraduates has a smaller chance,

$$\frac{300}{27,000} = \frac{1}{90}$$

Because we have two SRSs, it is easy to estimate opinions in the two groups separately. The quick and approximate method (page 46) tells us that the margin of error for a sample proportion will be about

$$\frac{1}{\sqrt{200}} = 0.07 \quad (\text{that is, } 7\%)$$

for grad students and about

$$\frac{1}{\sqrt{300}} = 0.058 \quad (\text{that is, } 5.8\%)$$

for undergraduates.

Because the sample in Example 8 deliberately overrepresents graduate students, the final analysis must adjust for this to get unbiased estimates of overall student opinion. Remember that our quick method works only for an SRS. In fact, a professional analysis would also take account of the fact that the population contains “only” 30,000 individuals—more job opportunities for statisticians.

4.2 A stratified sample. The statistics department at Cal Poly, San Luis Obispo, had 18 faculty members and 80 undergraduate majors in 2015. Use the *Simple Random Sample* applet, other software, or Table A, starting at line 111, to choose a stratified sample of one faculty member and one student to attend a reception being held by the university president.

**NOW IT'S
YOUR TURN**

EXAMPLE 9 The woes of telephone samples

In principle, it would seem that a telephone survey that dials numbers at random could be based on an SRS. Telephone surveys have little need for clustering. Stratifying can still reduce variability, however, and so telephone surveys often take samples in two stages: a stratified sample of telephone number prefixes (area code plus first three digits) followed by individual numbers (last four digits) dialed at random in each prefix.

The real problem with an SRS of telephone numbers is that too few numbers lead to households. Blame technology. Fax machines, modems, and cell phones demand new phone numbers. Between 1988 and 2008, the number of households in the United States grew by 29%, but the number of possible residential phone numbers grew by more than 120%. Some analysts believe that, in the near future, we may have to



New York, New

York New York City,

they say, is bigger, richer, faster, ruder. Maybe there's something to that. The sample survey firm Zogby International says that as a national average, it takes 5 telephone calls to reach a live person. When calling to New York, it takes 12 calls. Survey firms assign their best interviewers to make calls to New York and often pay them bonuses to cope with the stress.

increase the number of digits for telephone numbers from 10 (including the area code) to 12. This will further exacerbate this problem. Telephone surveys now use “list-assisted samples” that check electronic telephone directories to eliminate prefixes that have no listed numbers before random sampling begins. Fewer calls are wasted, but anyone living where all numbers are unlisted is missed. Prefixes with no listed numbers are, therefore, separately sampled (stratification again), perhaps with a smaller sample size than if included in the list-assisted sample, to fill the gap.

The proliferation of cell phones has created additional problems for telephone samples. As of December 2013, about 41% of households had cell phones only. Random digit dialing using a machine is not allowed for cell phone numbers. Phone numbers assigned to cell phones are determined by the location of the cell phone company providing the service and need not coincide with the actual residence of the user. This makes it difficult to implement sophisticated methods of sampling such as stratified sampling by geographic location.

It may be that the woes of telephone sampling prompted the Gallup Organization, in recent years, to drop the phrase “random sampling” from the description of their survey methods at the end of most of their polls. This presumably prevents misinterpreting the results as coming from simple random samples. In the organization’s detailed description of survey methods used for the Gallup World Poll and the Gallup Well-Being Index (available online at the Gallup website), the samples are described as involving random sampling.

The challenge of Internet surveys

The Internet is having a profound effect on many things people do, and this includes surveys. Using the Internet to conduct “Web surveys” is becoming increasingly popular. Web surveys have several advantages over more traditional survey methods. It is possible to collect large amounts of survey data at lower costs than traditional methods allow. Anyone can put survey questions on dedicated sites offering free services; thus, large-scale data collection is available to almost every person with access to the Internet. Furthermore, Web surveys allow one to deliver multimedia survey content to respondents, opening up new realms of survey possibilities that would be extremely difficult to implement using traditional methods. Some argue that eventually Web surveys will replace traditional survey methods.

Although Web surveys are easy to do, they are not easy to do well. The reasons include many of the issues we have discussed in this chapter. Three major problems are voluntary response, undercoverage, and nonresponse.

Voluntary response appears in several forms. Some Web surveys invite visitors to a particular website to participate in a poll. Misterpoll.com is one such example. Visitors to this site can participate in several ongoing polls, create their own poll, and respond multiple times to the same poll. Other Web surveys solicit participation through announcements in newsgroups, email invitations, and banner ads on high-traffic sites. An example is a series of 10 polls conducted by Georgia Tech University's Graphic, Visualization, and Usability Center (GVU) in the 1990s.

Although misterpoll.com indicates that the surveys on the site are primarily intended for entertainment, the Gvu polls appear to claim some measure of legitimacy. The website www.cc.gatech.edu/gvu/user_surveys/ states that the information from these surveys "is valued as an independent, objective view of developing Web demographics, culture, user attitudes, and usage patterns."

A third and more sophisticated example of voluntary response occurs when the polling organization creates what it believes to be a representative panel consisting of volunteers and uses panel members as a sampling frame. A random sample is selected from this panel, and those selected are invited to participate in the poll. A very sophisticated version of this approach is used by the Harris Poll Online.

Web surveys, such as the Harris Poll Online, in which a random sample is selected from a well-defined sampling frame are reasonable when the sampling frame clearly represents some larger population or when interest is only in the members of the sampling frame. An example are Web surveys that use systematic sampling to select every n th visitor to a site and the target population is narrowly defined as visitors to the site. Another example are some Web surveys on college campuses. All students may be assigned email addresses and have Internet access. A list of these email addresses serves as the sampling frame, and a random sample is selected from this list. If the population of interest is all students at this particular college, these surveys can potentially yield very good results. Here is an example of this type of Web survey.

EXAMPLE 10 Doctors and placebos

A placebo is a dummy treatment like a salt pill that has no direct effect on a patient but may bring about a response because patients expect it to. Do academic physicians who maintain private practices sometimes give their patients placebos? A Web survey of doctors in internal medicine departments at Chicago-area medical schools was possible because almost all doctors had listed email addresses.

An email was sent to each doctor explaining the purpose of the study, promising anonymity, and giving an individual a Web link for response. Result: 45% of respondents said they sometimes use placebos in their clinical practice.

Several other Web survey methods have been employed to eliminate problems arising from voluntary response. One is to use the Web as one of many alternative ways to participate in the survey. The Bureau of Labor Statistics and the U.S. Census Bureau have used this method. Another method is to select random samples from panels, but instead of relying on volunteers to form the panels, members are recruited using random sampling (for example, random digit dialing). Telephone interviews can be used to collect background information, identify those with Internet access, and recruit eligible persons to the panel. If the target population is current users of the Internet, this method should also potentially yield reliable results. The Pew Research Center has employed this method.

Perhaps the most ambitious approach, and one that attempts to obtain a random sample from a more general population, is the following. Take a probability sample from the population of interest. Provide all those selected with the necessary equipment and tools to participate in subsequent Web surveys. This methodology is similar in spirit to that used for the Nielsen TV ratings. It was employed by one company, InterSurvey, several years ago, although InterSurvey is no longer in business.

Several challenges remain for those who employ Web surveys. Even though Internet and email use is growing (according to the 2012 *Statistical Abstract of the United States*, as of 2010, 80% of American adults aged 18 and older have Internet access at home or work, and 71% have Internet access at home), there is still the problem of undercoverage if Web surveys are used to draw conclusions about all American adults aged 18 and older. Weighting responses to correct for possible biases does not solve the problem because studies indicate that Internet users differ in many ways that traditional methods of weighting do not account for.

In addition, even if 100% of Americans had Internet access, there is no list of Internet users that we can use as a sampling frame, nor is there anything comparable to random digit dialing that can be used to draw random samples from the collection of all Internet users.

Finally, Web surveys often have very high rates of nonresponse. Methods that are used in phone and mail surveys to improve response rates can help, but they make Web surveys more expensive and difficult, offsetting some of their advantages.

STATISTICAL CONTROVERSIES

The Harris Online Poll

The Harris Poll Online has created an online research panel of more than 6 million volunteers. According to the

Harris Poll Online website, the “panel consists of a diverse cross-section of people residing in the United States, as well as in

over 200 countries around the world,” and “this multimillion member panel consists of potential respondents who have been recruited through online, telephone, mail, and in-person approaches to increase population coverage and enhance representativeness.” One can join the panel at join.harrispollonline.com.

When the Harris Poll Online conducts a survey, this panel serves as the

sampling frame. A probability sample is selected from it, and statistical methods are used to weight the responses. In particular, the Harris Poll Online uses propensity score weighting, a proprietary Harris Interactive technique, which is also applied (when applicable) to adjust for respondents’ likelihood of being online. They claim that “this procedure provides added assurance of accuracy and representativeness.”

Are you convinced that the Harris Poll Online provides accurate information about well-defined populations such as all American adults? Why or why not?

For more information about the Harris Poll Online, visit

www.theharrispoll.com

There is more information in a special issue of *Public Opinion Quarterly*, available online at

<http://poq.oxfordjournals.org/content/72/5.toc>



Blend Images/Hill Street Studios/Getty Images

Probability samples

It's clear from Examples 7, 8, and 9, and from the challenges of using the Internet to conduct surveys, that designing samples is a business for experts. Even most statisticians don't qualify. We won't worry about such details. The big idea is that good sample designs use chance to select individuals from the population. That is, all good samples are *probability samples*.

Probability sample

A **probability sample** is a sample chosen by chance. We must know what samples are possible and what chance, or probability, each possible sample has.

Some probability samples, such as stratified samples, don't allow all possible samples from the population and may not give an equal chance to all the samples they do allow. As such, not all probability samples are random samples.

A stratified sample of 300 undergraduate students and 200 graduate students, for example, allows only samples with exactly that makeup. An SRS would allow any 500 students. Both are probability samples. We need only know that estimates from any probability sample share the nice properties of estimates from an SRS. Confidence statements can be made without bias and have smaller margins of error as the size of the sample increases. Nonprobability samples such as voluntary response samples do not share these advantages and cannot give trustworthy information about a population. Now that we know that most nationwide samples are more complicated than an SRS, we will usually go back to acting as if good samples were SRSs. That keeps the big idea and hides the messy details.

Questions to ask before you believe a poll

Opinion polls and other sample surveys can produce accurate and useful information if the pollster uses good statistical techniques and also works hard at preparing a sampling frame, wording questions, and reducing non-response. Many surveys, however, especially those designed to influence public opinion rather than just record it, do not produce accurate or useful information. Here are some questions to ask before you pay much attention to poll results.

- **Who carried out the survey?** Even a political party should hire a professional sample survey firm whose reputation demands that it follow good survey practices.
- **What was the population?** That is, whose opinions were being sought?
- **How was the sample selected?** Look for mention of random sampling.
- **How large was the sample?** Even better, find out both the sample size and the **margin of error** within which the results of 95% of all samples drawn as this one was would fall.
- **What was the response rate?** That is, what percentage of the original subjects actually provided information?
- **How were the subjects contacted?** By telephone? Mail? Face-to-face interview?
- **When was the survey conducted?** Was it just after some event that might have influenced opinion?
- **What were the exact questions asked?**

Academic survey centers and government statistical offices answer these questions when they announce the results of a sample survey. National opinion polls usually don't announce their response rate (which is often low) but do give us the other information. Editors and newscasters have the bad habit of cutting out these dull facts and reporting only the sample results. Many sample surveys by interest groups and local newspapers and TV stations don't answer these questions because their polling methods are, in fact, unreliable. If a politician, an advertiser, or your local TV station announces the results of a poll without complete information, be skeptical.

STATISTICS IN SUMMARY

Chapter Specifics

- Sampling in the real world is complex. Even professional sample surveys don't give exactly correct information about the population.
- There are many potential sources of error in sampling. The margin of error announced by a sample survey covers only **random sampling error**, the variation due to chance in choosing a random sample.
- Other types of error are in addition to the margin of error and can't be directly measured. **Sampling errors** come from the act of choosing a sample. Random sampling error and **undercoverage** are common types of sampling error. Undercoverage occurs when some members of the population are left out of the **sampling frame**, the list from which the sample is actually chosen.
- The most serious errors in most careful surveys, however, are **nonsampling errors**. These have nothing to do with choosing a sample—they are present even in a census.
- The single biggest problem for sample surveys is **nonresponse**: subjects can't be contacted or refuse to answer.
- Mistakes in handling the data (**processing errors**) and incorrect answers by respondents (**response errors**) are other examples of nonsampling errors.
- Finally, the exact **wording of questions** has a big influence on the answers.
- People who design sample surveys use statistical techniques that help correct nonsampling errors, and they also use **probability samples** more complex than simple random samples, such as **stratified samples**.

- You can assess the quality of a sample survey quite well by just looking at the basics: use of random samples, sample size and margin of error, the rate of nonresponse, and the wording of the questions.



In Chapter 3, we saw that random samples can provide a sound basis for drawing conclusions about a population parameter. In this chapter, we learned that even when we take a random sample, our conclusions can be weakened by undercoverage, processing errors, response errors, nonresponse, and wording of questions. We must pay careful attention to every aspect of how we collect data to ensure that the conclusions we make are valid. In some cases, more complex probability samples, such as stratified samples, can help correct nonsampling errors. This chapter provides a list of questions you can ask to help you assess the quality of the results of samples collected by someone else.

CASE STUDY Use what you have learned in this chapter to evaluate the Case **EVALUATED** Study that opened the chapter. In particular, do the following.

1. Answer the questions given in the section “Questions to Ask before You Believe a Poll” on page 80.
2. Are the results of the Pew poll useless? You may want to refer to the discussion on pages 76–79.



LaunchPad Online Resources
macmillan learning

- The video technology manuals explain how to select an SRS using Excel, JMP, Minitab, and the TI 83/84.
- The Statistical Applet *Simple Random Sample* can be used to select a simple random sample when the number of labels is 144 or fewer.

CHECK THE BASICS

For Exercise 4.1, see page 72; for Exercise 4.2, see page 75.

4.3 What does the margin of error include? When a margin of error is reported for a survey, it includes

(a) random sampling error and other practical difficulties like undercoverage and nonresponse.

(b) random sampling error, but not other practical difficulties like undercoverage and nonresponse.

(c) practical difficulties like undercoverage and nonresponse, but not random sampling error.

(d) None of the above is correct.

4.4 What kind of sample? Archaeologists plan to examine a sample of two-meter-square plots near an ancient Greek city for artifacts visible in the ground. They choose separate samples of plots from floodplain, coast, foothills, and high hills. What kind of sample is this?

- (a) A simple random sample
- (b) A voluntary response sample
- (c) A stratified sample
- (d) A cluster sample

4.5 Sampling issues. A sample of households in a community is selected at random from the telephone directory. In this community, 4% of households have no telephone, 10% have only cell phones, and another 25% have unlisted telephone numbers. The sample will certainly suffer from

- (a) nonresponse.
- (b) undercoverage.
- (c) false responses.
- (d) all of the above.

4.6 Question wording. Which of the following represents wording that will most likely *not* influence the answers?

- (a) Do you think that all instances of academic misconduct should be reported to the dean?

(b) Academic misconduct undermines the integrity of the university and education in general. Do you believe that all instances of academic misconduct should be reported to the dean?

(c) Academic misconduct can range from something as minor as using one's own work in two courses to major issues like cheating on exams and plagiarizing. Do you believe that all instances of academic misconduct should be reported to the dean?

(d) None of the above will influence the answers.

4.7 Sampling considerations. A statistics class has 10 graduate students and 40 undergraduate students. You want to randomly sample 10% of the students in the class. One graduate student and four undergraduate students are selected at random. Which of the following is *not* correct?

- (a) Because each student has a 10% chance of being selected, this is a simple random sample.
- (b) Because each sample includes exactly one graduate student and four undergraduate students, this is not a random sample.
- (c) It is possible to get a sample that contains only graduate students.
- (d) It is possible to get a sample that contains only undergraduate students.

CHAPTER 4 EXERCISES

4.8 What kind of error? Which of the following are sources of *sampling error* and which are sources of *nonsampling error*? Explain your answers.

- (a) The subject lies about past drug use.
- (b) A typing error is made in recording the data.

(c) Data are gathered by asking people to go to a website and answer questions online.

4.9 What kind of error? Each of the following is a source of error in a sample survey. Label each as *sampling error* or *nonsampling error*, and explain your answers.

- (a) The telephone directory is used as a sampling frame.
- (b) The subject cannot be contacted in five calls.
- (c) Interviewers choose people on the street to interview.

4.10 Not in the margin of error. According to a February 2008 *USA Today*/Gallup Poll, 43% of Americans identify themselves as baseball fans. That is low by recent standards, as an average of 49% of Americans have said they were fans of the sport since Gallup started tracking this measure in 1993. The high point came in 1998, when Sammy Sosa and Mark McGwire pursued (and ultimately surpassed) Roger Maris's single-season home run record, at which time 56% of Americans considered themselves baseball fans. The Gallup press release says:

For results based on this sample, one can say with 95% confidence that the maximum error attributable to sampling and other random effects is ± 5 percentage points.

Give one example of a source of error in the poll result that is *not* included in this margin of error.



4.11 Not in the margin of error. According to a June 2015 Gallup Poll, more

American adults ages 18 to 29 report being single and never married—the percentage has risen from 52% in 2004 to 64% in 2014. For American adults aged 30 to 39, this same percentage has risen modestly from 15% in 2004 to 19% in 2014. The Gallup press release says:

For results based on the total sample of Americans ages 18 to 29 or 30 to 39 in any given year, the maximum margin of sampling error is ± 3 percentage points at the 95% confidence level.

Give one example of a source of error in the poll result that is *not* included in this margin of error.

4.12 College parents. An online survey of college parents was conducted during February and March 2007. An email was sent to 41,000 parents who were listed in either the College Parents of America database or the Student Advantage database. Parents were invited to participate in the online survey. Out of those invited, 1727 completed the online survey. The survey protected the anonymity of those participating in the survey but did not allow more than one response from an individual IP address.

One of the survey results was that 33% of mothers communicate at least once a day with their child while at school.

(a) What was the *response rate* for this survey? (The response rate is the percentage of the planned sample—that is, those invited to participate—who responded.)

(b) Use the quick method (page 46) to estimate the margin of error for a random sample of size 1727.

(c) Do you think that the margin of error is a good measure of the accuracy of the survey's results? Explain your answer.

4.13 Polling customers. An online store chooses an SRS of 100 customers from its list of all people who have bought something from the store in the last year. It asks those selected how satisfied they are with the store's website. If it selected two SRSs of 100 customers at the same time, the two samples would give somewhat different results. Is this variation a source of sampling error or of nonsampling error? Would the survey's announced margin of error take this source of error into account?

4.14 Ring-no-answer. A common form of nonresponse in telephone surveys is "ring-no-answer." That is, a call is made to an active number but no one answers. The Italian National Statistical Institute looked at nonresponse to a government survey of households in Italy during the periods January 1 to Easter and July 1 to August 31. All calls were made between 7 and 10 P.M., but 21.4% gave "ring-no-answer" in one period versus 41.5% "ring-no-answer" in the other period. Which period do you think had the higher rate of no answers? Why? Explain why a high rate of nonresponse makes sample results less reliable.



4.15 Race relations. Here are two opinion poll questions asked about race relations in the United States.

We'd like to know how you would rate relations between various

groups in the United States these days. Would you say relations between whites and blacks are very good, somewhat good, somewhat bad, or very bad?

Do you think race relations in the U.S. are generally good or generally bad?

In response to the first question, 72% of non-Hispanic whites and 66% of blacks answered relations between blacks and whites are "very good" or "somewhat good." Sixty-one percent (61%) of those answering the second question responded "generally bad."

The first question came from a poll that was conducted in March 2015. The second question came from a poll that was conducted between April 30 and May 3 of that same year. Between the two polls, a man named Freddie Gray died after being in the custody of the Baltimore police. In what ways do you think this event may have affected the responses to the two different polls? Do you think the results would be different if the question from the second poll had been worded like the question from the first poll?

4.16 The environment and the economy. Here are two opinion poll questions asked in December 2009 about protecting the environment versus protecting the economy.

Often there are trade-offs or sacrifices people must make in deciding what is important to them. Generally speaking, when a trade-off has to be made, which is more important to you: stimulating the economy or protecting the environment?

Which worries you more: that the U.S. will NOT take the actions necessary to prevent the catastrophic effects of global warming because of fears those actions would harm the economy or that the U.S. WILL take actions to protect against global warming and those actions will cripple the U.S. economy?

In response to the first question, 61% said stimulating the economy was more important. But only 46% of those asked the second question said they were afraid that the United States will take actions to protect against global warming and that those actions will cripple the U.S. economy. Why do you think the second wording discouraged more people from expressing more concern about the economy than about the environment?

4.17 Amending the Constitution.

You are writing an opinion poll question about a proposed amendment to the Constitution. You can ask if people are in favor of “changing the Constitution” or “adding to the Constitution” by approving the amendment.

- (a) Why do you think the responses to these two questions will produce different percentages in favor?
- (b) One of these choices of wording will produce a much higher percentage in favor. Which one? Why?



4.18 Right to refuse services. The Supreme Court of the United States released a ruling on same-sex marriage in June 2015. Prior to the release of this ruling, many agencies

conducted polls about same-sex marriage. In April 2015, a Quinnipiac University poll asked two questions about businesses and services to gays and lesbians. Here are the two questions:

Do you think businesses should or should not be allowed to refuse services to gays and lesbians?

What if the business says homosexuality violates its owners' religious beliefs? In that case, do you think the business should or should not be allowed to refuse services to gays and lesbians?

One of these questions drew 35% saying the businesses should be allowed to refuse services; the other drew 26% with the same response. Which wording produced the higher percentage? Why?

4.19 Wording survey questions. Comment on each of the following as a potential sample survey question. Is the question clear? Is it slanted toward a desired response? (Survey questions on issues that one might regard as inflammatory are often prone to slanted wording.)

- (a) Which of the following best represents your opinion on gun control?
 1. The government should take away our guns.
 2. We have the right to keep and bear arms.
- (b) In light of skyrocketing gasoline prices, we should consider opening up a very small amount of Alaskan wilderness for oil exploration as a way of reducing our dependence on foreign oil. Do you agree or disagree?

(c) Do you think that the excessive restrictions placed on U.S. law enforcement agencies hampered their ability to detect the 9/11 terrorist plot before it occurred?

(d) Do you use drugs?

4.20 Bad survey questions. Write your own examples of bad sample survey questions.

(a) Write a biased question designed to get one answer rather than another.

(b) Write a question that is confusing so that it is hard to answer.



4.21 Appraising a poll. In January 2015, *The Wall Street Journal* published an article on satisfaction with the U.S. economy and other issues. The article noted 45% of Americans were very or somewhat satisfied with the state of the U.S. economy, even though 49% felt America is in a state of decline. News articles tend to be brief in describing sample surveys. Here is part of *The Wall Street Journal's* description of this poll:

The . . . poll was based on nationwide telephone interviews of 800 adults, including 280 respondents who use only a cellphone. . . . Individuals were selected proportionate to the nation's population in accordance with a probability sample. . . . The data's margin of error is plus or minus 3.46 percentage points.

Page 80 lists several “questions to ask” about an opinion poll. What answers does *The Wall Street Journal* give to each of these questions?



4.22 Appraising a poll. A June 2015 *New York Times* article about the inequality of wealth and income discussed the results of a sample survey that found, for example, that 67% of those surveyed think the gap between the rich and the poor in the United States is getting larger. Here is part of the *Times's* statement “How the Poll Was Conducted”:

The latest New York Times/CBS News Poll is based on telephone interviews conducted May 28 to 31 with 1,022 adults throughout the United States...

The sample of land-line telephone exchanges called was randomly selected by a computer...

The *Times's* statement goes on to say that cell phone numbers were also randomly selected by a computer and that the samples of landline and cell phone responses were combined and appropriately adjusted. Page 80 lists several “questions to ask” about an opinion poll. What answers does the *Times* give to each of these questions?

4.23 Closed versus open questions.

Two basic types of questions are closed questions and open questions. A closed question asks the subject for one or more of a fixed set of responses. An open question allows the subject to answer in his or her own words; the interviewer writes down the responses and classifies them later. An example of an open question is

What do you believe about the afterlife?

An example of a closed question is

What do you believe about the afterlife? Do you believe

- a. there is an afterlife and entrance depends only on your actions?*
- b. there is an afterlife and entrance depends only on your beliefs?*
- c. there is an afterlife and everyone lives there forever?*
- d. there is no afterlife?*
- e. I don't know.*

What are the advantages and disadvantages of open and closed questions?

4.24 Telling the truth? Many subjects don't give honest answers to questions about activities that are illegal or sensitive in some other way. One study divided a large group of white adults into thirds at random. All were asked if they had ever used cocaine. The first group was interviewed by telephone: 21% said Yes. In the group visited at home by an interviewer, 25% said Yes. The final group was interviewed at home but answered the question on an anonymous form that they sealed in an envelope. Of this group, 28% said they had used cocaine.

- (a) Which result do you think is closest to the truth? Why?
- (b) Give two other examples of behavior you think would be underreported in a telephone survey.

4.25 Did you vote? When the Current Population Survey asked the adults in its sample of 56,000 households if they voted in the 2012 presidential election, 61.8% said they had. The margin of error was less than 0.3%. In fact, only 58.2% of the adult population voted in that election. Why do

you think the CPS result missed by 12 times the margin of error?

4.26 A party poll. At a party, there are 20 students over age 21 and 40 students under age 21. You choose, at random, two of those over 21 and separately choose at random four of those under 21 to interview about attitudes toward alcohol. You have given every student at the party the same chance to be interviewed: what is that chance? Why is your sample not an SRS?

4.27 A stratified sample. A club has 30 student members and 10 faculty members. The students are

Aguirre	Gonzales	Rodriguez
Butterfield	Grebe	Ryndak
Caporuscio	Kemp	Soria
Carlson	Kessler	Spiel
Chilson	Koepnick	Stankiewicz
Clement	Macha	Steele
Cooper	Makis	Tong
Dobbs	Palacios	White
Freeman	Peralta	Williams
Girard	Risser	Zhang

The faculty members are

Atchade	Murphy
Craigmile	Nair
Everson	Nguyen
Fink	Romero
Hansen	Turkmen

The club can send three students and two faculty members to a convention. It decides to choose those who will go by random selection.

- (a) Use the *Simple Random Sample* applet, other technology, or Table A

to choose a stratified random sample of three students and two faculty members.

(b) What is the chance that the student named White is chosen? What is the chance that faculty member Romero is chosen?

4.28 A stratified sample. A state university has 4900 in-state students and 2100 out-of-state students. A financial aid officer wants to poll the opinions of a random sample of students. In order to give adequate attention to the opinions of out-of-state students, the financial aid officer decides to choose a stratified random sample of 200 in-state students and 200 out-of-state students. The officer has alphabetized lists of in-state and out-of-state students.

(a) Explain how you would assign labels and use random digits to choose the desired sample. Use the *Simple Random Sample* applet, other technology, or Table A at line 122 and give the first five in-state students and the first five out-of-state students in your sample.

(b) What is the chance that any one of the 4900 in-state students will be in your sample? What is the chance that any one of the 2100 out-of-state students will be in your sample?

4.29 Sampling by accountants. Accountants use stratified samples during audits to verify a company's records of such things as accounts receivable. The stratification is based on the dollar amount of the item and often includes 100% sampling of the largest items. One company reports 5000 accounts receivable. Of these, 100 are in amounts over \$50,000; 500

are in amounts between \$1000 and \$50,000; and the remaining 4400 are in amounts under \$1000. Using these groups as strata, you decide to verify all of the largest accounts and to sample 5% of the midsize accounts and 1% of the small accounts. How would you label the two strata from which you will sample? Use the *Simple Random Sample* applet, other technology, or Table A, starting at line 115, to select *only the first five* accounts from each of these strata.

4.30 A sampling paradox? Example 8 compares two SRSs of a university's undergraduate and graduate students. The sample of undergraduates contains a smaller fraction of the population, 1 out of 90, versus 1 out of 15 for graduate students. Yet sampling 1 out of 90 undergraduates gives a smaller margin of error than sampling 1 out of 15 graduate students. Explain to someone who knows no statistics why this happens.



4.31 Appraising a poll. Exercise 4.22 gives part of the description of a sample survey from *The New York Times*. It appears that the sample was taken in several stages. Why can we say this? The first stage no doubt used a stratified sample, though the *Times* does not say this. Explain why it would be bad practice to use an SRS from all possible telephone numbers rather than a stratified sample of landline-only, cell-phone-only, and dual-phone users.



4.32 Multistage sampling. An article in the journal *Science* looks at differences in attitudes toward genetically modified foods between Europe and

the United States. This calls for sample surveys. The European survey chose a sample of 1000 adults in each of 17 European countries. Here's part of the description: "The Eurobarometer survey is a multistage, random-probability face-to-face sample survey."

- (a) What does "multistage" mean?
- (b) You can see that the first stage was stratified. What were the strata?
- (c) What does "random-probability sample" mean?

4.33 Online courses in high schools?

What do adults believe about requiring online courses in high schools? Are opinions different in urban, suburban, and rural areas? To find out, researchers wanted to ask adults this question:

It has become common for education courses after high school to be taken online. In your opinion, should public high schools in your community require every student to take at least one course online while in high school?

Because most people live in heavily populated urban and suburban areas, an SRS might contain few rural adults. Moreover, it is too expensive to choose people at random from a large region. We should start by choosing school districts rather than people. Describe a suitable sample design for this study, and explain the reasoning behind your choice of design.

4.34 Systematic random samples. The last stage of the Current Population Survey (Example 7) uses a **systematic random sample**. An example

will illustrate the idea of a systematic sample. Suppose that we must choose four rooms out of the 100 rooms in a dormitory. Because $100/4 = 25$, we can think of the list of 100 rooms as four lists of 25 rooms each. Choose one of the first 25 rooms at random, using Table A. The sample will contain this room and the rooms 25, 50, and 75 places down the list from it. If 13 is chosen, for example, then the systematic random sample consists of the rooms numbered 13, 38, 63, and 88. Use Table A to choose a systematic random sample of five rooms from a list of 200. Enter the table at line 120.

4.35 Systematic isn't simple. Exercise 4.34 describes a systematic random sample. Like an SRS, a systematic sample gives all individuals the same chance to be chosen. Explain why this is true, then explain carefully why a systematic sample is nonetheless *not* an SRS.

4.36 Planning a survey of students.

The student government plans to ask a random sample of students their opinions about on-campus parking. The university provides a list of the 20,000 enrolled students to serve as a sampling frame.

- (a) How would you choose an SRS of 200 students?
- (b) How would you choose a systematic sample of 200 students? (See Exercise 4.34 to learn about systematic samples.)
- (c) The list shows whether students live on campus (8000 students) or off campus (12,000 students). How would you choose a stratified sample of 50 on-campus students and 150 off-campus students?

4.37 Sampling students. You want to investigate the attitudes of students at your school toward the school's policy on extra fees for lab courses. You have a grant that will pay the costs of contacting about 500 students.

(a) Specify the exact population for your study. For example, will you include part-time students?

(b) Describe your sample design. For example, will you use a stratified sample with student majors as strata?

(c) Briefly discuss the practical difficulties that you anticipate. For example, how will you contact the students in your sample?

4.38 Mall interviews. Example 1 in Chapter 2 (page 23) describes mall interviewing. This is an example of a convenience sample. Why do mall interviews not produce probability samples?



4.39 Partial-birth abortion?

Here are three opinion poll questions on the same issue, with the poll results:

As you may know, the Supreme Court recently upheld a law that makes

the procedure commonly known as a partial birth abortion illegal. Do you favor or oppose this ruling by the Supreme Court? Result: 53% favor; 34% oppose.

[T]he Supreme Court . . . upheld a law that makes . . . partial birth abortion illegal. [This] procedure [is] performed in the late-term of pregnancy, when in some cases the baby is old enough to survive. . . . The court's ruling . . . [doesn't] make an exception for the [mother's] health. . . . Do you favor or oppose this ruling . . . ? Result: 47% favor; 43% oppose.

. . . I would like to ask your opinion about a specific abortion procedure known as a "late-term" . . . or "partial-birth" abortion, which is sometimes performed . . . during the last few months of pregnancy. Do you think . . . the government should make this procedure illegal, or do you think . . . the procedure should be legal? Result: 66% illegal; 28% legal.

Using this example, discuss the difficulty of using responses to opinion polls to understand public opinion.



EXPLORING THE WEB

Follow the QR code to access exercises.

