20.19 How many American adults aged 18 and over must be interviewed to estimate the proportion who own MP3 players with an error of ±0.02 with 99% confidence? Use 0.5 as the conservative guess for \( p \).
(a) \( n = 1692 \)  \hspace{1cm} (b) \( n = 2401 \)  \hspace{1cm} (c) \( n = 4148 \)

20.20 An opinion poll asks an SRS of 100 college seniors how they view their job prospects. In all, 53 say “Good.” The plus four 95% confidence interval for estimating the proportion of all college seniors who think their job prospects are good is (a) 0.529 ± 0.096.  \hspace{1cm} (b) 0.529 ± 0.098.  \hspace{1cm} (c) 0.529 ± 0.049.

20.21 The sample survey in Exercise 20.20 actually called 130 seniors, but 30 of the seniors refused to answer. This nonresponse could cause the survey result to be in error. The error due to nonresponse (a) is in addition to the margin of error found in Exercise 20.20. \hspace{1cm} (b) is included in the margin of error found in Exercise 20.20. \hspace{1cm} (c) can be ignored because it isn’t random.

20.22 Does the poll in Exercise 20.20 give reason to conclude that more than half of all seniors think their job prospects are good? The hypotheses for a test to answer this question are (a) \( H_0: p = 0.5, H_a: p > 0.5 \). \hspace{1cm} (b) \( H_0: p > 0.5, H_a: p = 0.5 \). \hspace{1cm} (c) \( H_0: p = 0.5, H_a: p \neq 0.5 \).

20.23 The value of the \( z \) statistic for the test of the previous exercise is about (a) \( z = 12 \). \hspace{1cm} (b) \( z = 6 \). \hspace{1cm} (c) \( z = 0.6 \).

20.24 A Gallup Poll found that only 28% of American adults expect to inherit money or valuable possessions from a relative. The poll’s margin of error was 3%. This means that (a) the poll used a method that gets an answer within 3% of the truth about the population 95% of the time. \hspace{1cm} (b) we can be sure that the percent of all adults who expect an inheritance is between 25% and 31%. \hspace{1cm} (c) if Gallup takes another poll using the same method, the results of the second poll will lie between 25% and 31%.

### Chapter 20 Exercises

We recommend using the plus four method for all confidence intervals for a proportion. However, the large-sample method is acceptable when the guidelines for its use are met.

20.25 Do smokers know that smoking is bad for them? The Harris Poll asked a sample of smokers, "Do you believe that smoking will probably shorten your life, or not?" Of the 1010 people in the sample, 848 said "Yes."

(a) Harris called residential telephone numbers at random in an attempt to contact an SRS of smokers. Based on what you know about national sample surveys, what is likely to be the biggest weakness in the survey? \hspace{1cm} (b) We will nonetheless act as if the people interviewed are an SRS of smokers. Give a 95% confidence interval for the percent of smokers who agree that smoking will probably shorten their lives.

20.26 Reporting cheating. Students are reluctant to report cheating by other students. A student project put this question to an SRS of 172 undergraduates at a large university: “You witness two students cheating on a quiz. Do you go to the professor?” Only 19 answered “Yes.” Give a 95% confidence interval for the proportion of all undergraduates at this university who would report cheating.

20.27 Harris announces a margin of error. Exercise 20.25 describes a Harris Poll survey of smokers in which 848 of a sample of 1010 smokers agreed that smoking would probably shorten their lives. Harris announces a margin of error of ±3 percentage points for all samples of about this size. Opinion polls announce the margin of error for 95% confidence.

(a) What is the actual margin of error (in percent) for the large-sample confidence interval from this sample? \hspace{1cm} (b) The margin of error is largest when \( \hat{p} = 0.5 \). What would the margin of error (in percent) be if the sample had resulted in \( \hat{p} = 0.5 \)? \hspace{1cm} (c) Why do you think that Harris announces a ±3% margin of error for all samples of about this size?

20.28 Prayer among the Millennials, continued. The Millennial generation (so called because they were born after 1980 and began to come of age around the year 2000) are less religiously active than older Americans. One of the questions in the General Social Survey in 2008 was “How often does the respondent pray?” Among the 385 respondents in the survey between 18 and 30 years of age, 247 prayed at least once a week.
20.29 Internet searches and cell phones. Pew Internet and American Life Project asked a random sample of 2485 cell phone users whether they had used their cell phone to look up health or medical information. Of these, 422 said "Yes."

(a) Pew dialed cell phone telephone numbers at random in the continental United States in an attempt to contact a random sample of adults. Based on what you know about national sample surveys, what is likely to be the biggest weakness in the survey?

(b) Act as if the sample is an SRS. Give a large-sample 90% confidence interval for the proportion $p$ of all cell phone users who have used their cell phone to look up health or medical information.

(c) Three out of the five most popular health-related searches on cell phones have to do with sex: "pregnancy," "herpes," and "STD" (sexually transmitted diseases). Sex-related queries don’t even show up on Google and Yahoo’s lists of the top five health searches on computers. What do you think explains the difference in the topics of health-related searches on cell phones versus computers? When drawing conclusions from a sample, you must always be careful to take into account the relevant population.

20.30 Which font? Plain type fonts such as Times New Roman are easier to read than fancy fonts such as Gigi. A group of 25 volunteer subjects read the same text in both fonts. (This is a matched pairs design. One-sample procedures for proportions, like those for means, are used to analyze data from matched pairs designs.) Of the 25 subjects, 17 said that they preferred Times New Roman for Web use. But 20 said that Gigi was more attractive.

(a) Because the subjects were volunteers, conclusions from this sample can be challenged. Show that the sample size condition for the large-sample confidence interval is not met, but that the condition for the plus four interval is met.

(b) Give a 95% confidence interval for the proportion of all adults who prefer Times New Roman for Web use. Give a 90% confidence interval for the proportion of all adults who think Gigi is more attractive.

20.31 Testing the waters. In August 2010, the Columbus Dispatch took water samples at 20 Ohio State Park swimming areas and tested for fecal coliform, which are bacteria found in human and animal feces. Experts warn that the tests are a snapshot of the quality of the water at the time they were taken, and levels can change as weather and other conditions vary. An unsafe level of fecal coliform means that there’s a higher chance that disease-causing bacteria are present and more risk that a swimmer will become ill. Of the 20 swimming areas tested, 13 were found to have unsafe levels of fecal coliform according to state standards. Assume that the swimming areas tested represent a random sample of swimming areas throughout the state.

(a) Show that the conditions for the large-sample confidence interval are not met. Show that the conditions for the plus four interval are met.

(b) Use the plus four method to give a 90% confidence interval for the percent of Ohio State Park swimming areas that have unsafe levels of fecal coliform.

20.32 Running red lights. A random digit dialing telephone survey of 880 drivers asked, “Recalling the last ten traffic lights you drove through, how many of them were red when you entered the intersections?” Of the 880 respondents, 171 admitted that at least one light had been red.

(a) Give a 95% confidence interval for the proportion of all drivers who ran one or more of the last ten red lights they encountered.

(b) Nonresponse is a practical problem for this survey—only 21.6% of calls that reached a live person were completed. Another practical problem is that people may not give truthful answers. What is the likely direction of the bias: do you think more or fewer than 171 of the 880 respondents really ran a red light? Why?

20.33 The IRS plans an SRS. The Internal Revenue Service plans to examine an SRS of individual federal income tax returns from each state. One variable of interest is the proportion of returns claiming itemized deductions. The total number of tax returns in a state varies from more than 15 million in California to fewer than 250,000 in Wyoming.

(a) Will the margin of error for estimating the population proportion change from state to state if an SRS of 2000 tax returns is selected in each state? Explain your answer.

(b) Will the margin of error change from state to state if an SRS of 1% of all tax returns is selected in each state? Explain your answer.

20.34 Customer satisfaction. An automobile manufacturer would like to know what proportion of its customers are not satisfied with the service provided by the local dealer.
The customer relations department will survey a random sample of customers and compute a 99% confidence interval for the proportion who are not satisfied.

(a) Past studies suggest that this proportion will be about 0.2. Find the sample size needed if the margin of error of the confidence interval is to be about 0.015.

(b) When the sample of the size found in (a) is actually contacted, 10% of the sample say they are not satisfied. What is the margin of error of the 99% confidence interval?

**20.35 Surveying students.** You are planning a survey of students at a large university to determine what proportion favor an increase in student fees to support an expansion of the student newspaper. Using records provided by the registrar, you can select a random sample of students. You will ask each student in the sample whether he or she is in favor of the proposed increase. Your budget will allow a sample of 100 students.

(a) For a sample of size 100, construct a table of the margins of error for 95% confidence intervals when \( \hat{p} \) takes the values 0.1, 0.3, 0.5, 0.7, and 0.9.

(b) A former editor of the student newspaper offers to provide funds for a sample of size 500. Repeat the margin of error calculations in (a) for the larger sample size. Then write a short thank-you note to the former editor describing how the larger sample size will improve the results of the survey.

In responding to Exercises 20.36 to 20.44, follow the Plan, Solve, and Conclude steps of the four-step process.

**20.36 College-educated parents.** The National Assessment of Educational Progress (NAEP) includes a “long-term trend” study that tracks reading and mathematics skills over time and obtains demographic information. In the 2008 study, a random sample of 9600 17-year-old students was selected.25 The NAEP sample used a multistage design, but the overall effect is quite similar to an SRS of 17-year-olds who are in school.

(a) In the sample, 46% of students had at least one parent who was a college graduate. Estimate with 99% confidence the proportion of all 17-year-olds in 2008 who had at least one parent who was a college graduate.

(b) The sample does not include 17-year-olds who dropped out of school, so your estimate is valid only for students. Do you think that the proportion of all 17-year-olds with at least one parent who was a college graduate would be higher or lower than 46%? Explain.

**20.37 Shrubs that survive fires.** Some shrubs have the useful ability to resprout from their roots after their tops are destroyed. Fire is a particular threat to shrubs in dry climates, as it can injure the roots as well as destroy the aboveground material. One study of resprouting took place in a dry area of Mexico.26 The investigators clipped the tops of samples of several species of shrubs. In some cases, they also applied a propane torch to the stumps to simulate a fire. Of 12 specimens of the shrub Krameria cyritosides, 5 resprouted after fire. Estimate with 90% confidence the proportion of all shrubs of this species that will resprout after fire.

**20.38 Downloading music.** A husband and wife, Ted and Suzanne, share a digital music player that has a feature that randomly selects which song to play. A total of 3476 songs have been loaded into the player, some by Ted and the rest by Suzanne. They are interested in determining whether they have each loaded a different proportion of songs into the player. Suppose that when the player was in the random-selection mode, 22 of the first 30 songs selected were songs loaded by Suzanne. Let \( \hat{p} \) denote the proportion of songs that were loaded by Suzanne. State the null and alternative hypotheses to be tested. How strong is the evidence that Ted and Suzanne have each loaded a different proportion of songs into the player?

**20.39 Opinions about evolution.** A sample survey funded by the National Science Foundation asked a random sample of American adults about biological evolution.27 One question asked subjects to answer “True,” “False,” or “Not sure” to the statement “Human beings, as we know them today, developed from earlier species of animals.” Of the 1484 respondents, 594 said “True.” What can you say with 95% confidence about the percent of all American adults who think that humans developed from earlier species of animals?

**20.40 Order in choice.** Does the order in which wine is presented make a difference? Several choices of wine are presented one at a time, and the subject is then asked to choose his or her preferred wine at the end of the sequence. In one study, subjects were asked to taste two wine samples in sequence. Both samples given to a subject were the same wine, although the subjects were expecting to taste two different samples of a particular variety. Of the 32 subjects in the study, 22 selected the wine presented first when presented with two identical wine samples.28

(a) Give a 95% confidence interval for the proportion of subjects who would select the first choice presented.

(b) The subjects were recruited in Ontario, Canada, via advertisements to participate in a study of “attitudes and values towards wine.” What assumption are you making about these subjects?
20.41 Opinions about evolution, continued. Does the sample in Exercise 20.39 give good evidence to support the claim “Fewer than half of American adults think that humans developed from earlier species of animals”?

20.42 Order in choice, continued. Do the data in Exercise 20.40 give good reason to conclude that the subjects are not equally likely to choose either of the two wines when presented with two identical wine samples in sequence? Can we generalize our conclusions to all wine tasters? Explain.

20.43 Chick-fil-A gets it right. Which fast-food chain fills orders most accurately at the drive-thru window? The Quick Service Restaurant (QSR) magazine drive-thru study involved a total of 7594 visits to restaurants in the 25 largest fast-food chains in all 50 states. All visits occurred during the lunch hours of 11:00 A.M. to 2:30 P.M. or during the dinner hours of 4:00 to 7:00 P.M. During each visit, the researcher ordered a main item, a side item, and a drink. One item was left off of each order; for example, a field researcher could order a burger with no pickles. After receiving the order, all food and drink items were checked for complete accuracy. Any food or drink item received that was not exactly as ordered resulted in the order being classified as inaccurate. Also included in the measurement of accuracy were condiments asked for, napkins, straws, and correct change. Any errors in these resulted in the order being classified as inaccurate. Chick-fil-A had the fewest inaccuracies, with only 14 of 196 orders classified as inaccurate. What proportion of orders are filled accurately by Chick-fil-A? (Use 95% confidence.)

20.44 Order in choice: planning a study. How large a sample would be needed to obtain margin of error ±0.05 in the study of choice order for tasting wine? Use the \( \hat{p} \) from Exercise 20.40 as your guess for the unknown \( p \).

EXPLORING THE WEB

20.45 Health care access/coverage. The Behavioral Risk Factor Surveillance System (BRFSS) is an ongoing data collection program designed to measure behavioral risk factors for the adult population (18 years of age or older) living in households. Data are collected from a random sample of adults (one per household) through a telephone survey. Go to the Web site apps.nccd.cdc.gov/BRFSS/ and under “Category” go to “Health Care Access/Coverage.” Under the topic “Adults aged 18-64 who have any kind of health care coverage,” you will find the percent with coverage in each state.

(a) Which state has the highest percent of coverage, and what is the reported value? Which state has the lowest percent, and what is its value? Are the reported percents statistics or parameters?

(b) Choose a state of interest to you and click on the link. In the table that opens, there is a line for \( n \), and the entries are the numbers who answered “Yes” and “No.” Find the percent in the sample who answered “Yes.” Notice that it is different from the percent reported in the table. The table estimates are weighted to try to reduce bias. If it is determined that certain portions of the population are underrepresented in the sample, then
that portion of the sample receives more weight when computing the estimate of the percent. The assumptions for an SRS are rarely met in practice, and more complicated methods are often necessary to estimate proportions and compute confidence intervals.

**20.46 Find a poll.** Search the Web for a recent poll in which the sample statistic is a proportion, for example, the proportion in the sample responding “Yes” to a question. Calculate a 95% confidence interval for the population proportion (assume that the sample is a random sample). State the question asked, how the sample was collected, the sample size, and the population of interest. Possible Web sites are [www.gallup.com](http://www.gallup.com) and [www.cbsnews.com/sections/opinion/polls/main500160.shtml](http://www.cbsnews.com/sections/opinion/polls/main500160.shtml).