

SECTION 8.1 EXERCISES

- 8.1 To profitably produce a planned upgrade of a software product you make you must charge customers \$100. Are your customers willing to pay this much? You contact a random sample of 40 customers and find that 11 would pay \$100 for the upgrade. Find a 95% confidence interval for the proportion of all of your customers (the population) who would be willing to buy the upgrade for \$100.
- 8.2 A poll of 811 adults aged 18 or older asked about purchases that they intended to make for the upcoming holiday season.⁶ One of the questions asked about what kind of gift they intended to buy for the person on whom they would spend the most. Clothing was the first choice of 487 people. Give a 99% confidence interval for the proportion of people in this population who intend to buy clothing as their first choice.
- 8.3 Yesterday, your top salesperson called on 5 customers and obtained orders for your new product from all 5. Suppose that it is reasonable to view these 5 customers as a random sample of all of her customers.
- Give the Wilson estimate of the proportion of her customers who will buy the new product. Notice that we don't estimate that all customers will buy, even though all 5 in the sample did.
 - Give the margin of error for 95% confidence. (You may see that the upper endpoint of the confidence interval is greater than 1. In that case, take the upper endpoint to be 1.)
 - Do the results apply to all of your sales force? Explain why or why not.
- 8.4 When trying to hire managers and executives, companies sometimes verify the academic credentials described by the applicants. One company that performs these checks summarized its findings for a six-month period. Of the 84 applicants whose credentials were checked, 15 lied about having a degree.⁷
- Find the proportion of applicants who lied about having a degree and the standard error.
 - Consider these data to be a random sample of credentials from a large collection of similar applicants. Give a 95% confidence interval for the true proportion of applicants who lie about having a degree.
- 8.5 Refer to the previous exercise. Suppose that 10 applicants lied about their major. Can we conclude that a total of $25 = 15 + 10$ applicants lied about having a degree or about their major? Explain your answer.
- 8.6 In the United States approximately 900 people die in bicycle accidents each year. One study examined the records of 1711 bicyclists aged 15 or older who were fatally injured in bicycle accidents between 1987 and 1991 and were tested for alcohol. Of these, 542 tested positive for alcohol (blood alcohol concentration of 0.01% or higher).⁸

- (a) Summarize the data with appropriate descriptive statistics.
- (b) To do statistical inference for these data, we think in terms of a model where p is a parameter that represents the probability that a tested bicycle rider is positive for alcohol. Find a 99% confidence interval for p .
- (c) Can you conclude from your statistical analysis of this study that alcohol causes fatal bicycle accidents?

8.7 Refer to the previous exercise. In this study 386 bicyclists had blood alcohol levels above 0.10%, a level defining legally drunk in many states. Give a 99% confidence interval for the proportion who were legally drunk according to this criterion.

8.8 One question in the Christmas tree market survey described in Examples 8.6 and 8.7 (page 580) was “Did you have a Christmas tree last year?” Of the 500 respondents, 421 answered “Yes.”

- (a) Find the sample proportion and its standard error.
- (b) Give a 90% confidence interval for the proportion of Indiana households who had a Christmas tree last year.

8.9 As part of a quality improvement program, your mail-order company is studying the process of filling customer orders. According to company standards, an order is shipped on time if it is sent within 3 working days of the time it is received. You select an SRS of 200 of the 5000 orders received in the past month for an audit. The audit reveals that 185 of these orders were shipped on time. Find a 95% confidence interval for the true proportion of the month’s orders that were shipped on time.

8.10 Large trees growing near power lines can cause power failures during storms when their branches fall on the lines. Power companies spend a great deal of time and money trimming and removing trees to prevent this problem. Researchers are developing hormone and chemical treatments that will stunt or slow tree growth. If the treatment is too severe, however, the tree will die. In one series of laboratory experiments on 216 sycamore trees, 41 trees died. Give a 95% confidence interval for the proportion of sycamore trees that would be expected to die from this particular treatment.

8.11 In recent years over 70% of first-year college students responding to a national survey have identified “being well-off financially” as an important personal goal. A state university finds that 103 of an SRS of 150 of its first-year students say that this goal is important. Give a 95% confidence interval for the proportion of all first-year students at the university who would identify being well-off as an important personal goal.

8.12 In Exercise 8.1 we found that 11 customers from a random sample of 40 would be willing to buy a software upgrade that costs \$100. If the upgrade is to be profitable, you will need to sell it to more than 20% of your

- customers. Do the sample data give good evidence that more than 20% are willing to buy?
- Formulate this problem as a hypothesis test. Give the null and alternative hypotheses. Will you use a one-sided or a two-sided alternative? Why?
 - Carry out the significance test. Report the test statistic and the P -value.
 - Should you proceed with plans to produce and market the upgrade?
- 8.13** In each of the following cases state whether or not the normal approximation to the binomial should be used for a significance test on a population proportion p .
- $n = 10$ and $H_0: p = 0.6$.
 - $n = 100$ and $H_0: p = 0.4$.
 - $n = 2000$ and $H_0: p = 0.996$.
 - $n = 500$ and $H_0: p = 0.25$.
- 8.14** Refer to the Christmas tree market survey described in Examples 8.6 and 8.7 (page 580). Of the 500 responding households, 38% were from rural areas (including small towns), and the other 62% were from urban areas (including suburbs). According to the census, 36% of Indiana households are in rural areas, and the remaining 64% are in urban areas. Let p be the proportion of rural respondents. Set up hypotheses about p_0 and perform a test of significance to examine how well the sample represents the state in regard to rural versus urban residence. Summarize your results.
- 8.15** Refer to the previous exercise. There we arbitrarily chose to state the hypotheses in terms of the proportion of rural respondents. We could as easily have used the proportion of *urban* respondents.
- Write hypotheses in terms of the proportion of urban residents to examine how well the sample represents the state in regard to rural versus urban residence.
 - Perform the test of significance and summarize the results.
 - Compare your results with the results of the previous exercise. Summarize and generalize your conclusion.
- 8.16** A national opinion poll found that 44% of all American adults agree that parents should be given vouchers good for education at any public or private school of their choice. The result was based on a small sample. How large an SRS is required to obtain a margin of error of ± 0.03 (that is, $\pm 3\%$) in a 95% confidence interval? (Use the previous poll's result to obtain the guessed value p^* .)
- 8.17** An entomologist samples a field for egg masses of a harmful insect by placing a yard-square frame at random locations and carefully examining the ground within the frame. An SRS of 75 locations selected from a county's pastureland found egg masses in 13 locations. Give a 90% confidence interval for the proportion of all possible locations that are infested.

- 8.18** Of the 500 respondents in the Christmas tree market survey, 44% had no children at home and 56% had at least one child at home. The corresponding figures for the most recent census are 48% with no children and 52% with at least one child. Test the null hypothesis that the telephone survey technique has a probability of selecting a household with no children that is equal to the value obtained by the census. Give the z statistic and the P -value. What do you conclude?
- 8.19** The English mathematician John Kerrich tossed a coin 10,000 times and obtained 5067 heads.
- Is this significant evidence at the 5% level that the probability that Kerrich's coin comes up heads is not 0.5?
 - Use a 95% confidence interval to find the range of probabilities of heads that would not be rejected at the 5% level.
- 8.20** A matched pairs experiment compares the taste of instant versus fresh-brewed coffee. Each subject tastes two unmarked cups of coffee, one of each type, in random order and states which he or she prefers. Of the 50 subjects who participate in the study, 19 prefer the instant coffee. Let p be the probability that a randomly chosen subject prefers freshly brewed coffee to instant coffee. (In practical terms, p is the proportion of the population who prefer fresh-brewed coffee.)
- Test the claim that a majority of people prefer the taste of fresh-brewed coffee. Report the z statistic and its P -value. Is your result significant at the 5% level? What is your practical conclusion?
 - Find a 90% confidence interval for p .
- 8.21** Shereka, a starting player for a major college basketball team, made only 36.2% of her free throws last season. During the summer she worked on developing a softer shot in the hope of improving her free-throw accuracy. In the first eight games of this season Shereka made 22 free throws in 42 attempts. Let p be her probability of making each free throw she shoots this season.
- State the null hypothesis H_0 that Shereka's free-throw probability has remained the same as last year and the alternative H_a that her work in the summer resulted in a higher probability of success.
 - Calculate the z statistic for testing H_0 versus H_a .
 - Do you accept or reject H_0 for $\alpha = 0.05$? Find the P -value.
 - Give a 90% confidence interval for Shereka's free-throw success probability for the new season. Are you convinced that she is now a better free-throw shooter than last season?
 - What assumptions are needed for the validity of the test and confidence interval calculations that you performed?
- 8.22** One of your employees has suggested that your company develop a new product. You decide to take a random sample of your customers and ask whether or not there is interest in the new product. The response is on a

- 1 to 5 scale with 1 indicating “definitely would not purchase”; 2, “probably would not purchase”; 3, “not sure”; 4, “probably would purchase”; and 5 “definitely would purchase.” For an initial analysis, you will record the responses 1, 2, and 3 as “No” and 4 and 5 as “Yes”. What sample size would you use if you wanted the 95% margin of error to be 0.10 or less?
- 8.23** Refer to the previous exercise. Suppose that after reviewing the results of the previous survey, you proceeded with preliminary development of the product. Now you are at the stage where you need to decide whether or not to make a major investment to produce and market it. You will use another random sample of your customers but now you want the margin of error to be smaller. What sample size would you use if you wanted the 95% margin of error to be 0.05 or less?
- 8.24** You want to estimate the proportion of students at your college or university who are employed for 10 or more hours per week while classes are in session. You plan to present your results by a 95% confidence interval. Using the guessed value $p^* = 0.40$, find the sample size required if the interval is to have an approximate margin of error of $m = 0.06$.
- 8.25** Land’s Beginning is a company that sells its merchandise through the mail. It is considering buying a list of addresses from a magazine. The magazine claims that at least 25% of its subscribers have high incomes (they define this to be household income in excess of \$100,000). Land’s Beginning would like to estimate the proportion of high-income people on the list. Checking income is very difficult and expensive but another company offers this service. Land’s Beginning will pay to find incomes for an SRS of people on the magazine’s list. They would like the margin of error of the 95% confidence interval for the proportion to be 0.05 or less. Use the guessed value $p^* = 0.25$ to find the required sample size.
- 8.26** Refer to the previous exercise. For each of the following variations on the design specifications, state whether the required sample size will be higher, lower, or the same as that found above.
- Use a 90% confidence interval.
 - Change the allowable margin of error to 0.10.
 - Use a planning value of $p^* = 0.30$.
 - Use a different company to do the income checks.
- 8.27** A student organization wants to start a nightclub for students under the age of 21. To assess support for this proposal, they will select an SRS of students and ask each respondent if he or she would patronize this type of establishment. They expect that about 60% of the student body would respond favorably. What sample size is required to obtain a 95% confidence interval with an approximate margin of error of 0.08? Suppose that 50% of the sample responds favorably. Calculate the margin of error of the confidence interval.

- 8.28** An automobile manufacturer would like to know what proportion of its customers are dissatisfied with the service received from their local dealer. The customer relations department will survey a random sample of customers and compute a 95% confidence interval for the proportion that are dissatisfied. From past studies, they believe that this proportion will be about 0.25. Find the sample size needed if the margin of error of the confidence interval is to be about 0.02. Suppose 15% of the sample say that they are dissatisfied. What is the margin of error of the 95% confidence interval?
- 8.29** You have been asked to survey students at a large college to determine the proportion who favor an increase in student fees to support an expansion of the student newspaper. Each student will be asked whether he or she is in favor of the proposed increase. Using records provided by the registrar you can select a random sample of students from the college. After careful consideration of your resources, you decide that it is reasonable to conduct a study with a sample of 10 students. For this sample size, construct a table of the margins of error for 95% confidence intervals when \hat{p} takes the values 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9.
- 8.30** A former editor of the student newspaper agrees to underwrite the study in the previous exercise because she believes the results will demonstrate that most students support an increase in fees. She is willing to provide funds for a sample of size 400. Write a short summary for your benefactor of why the increased sample size will provide better results.

8.2 Comparing Two Proportions

Because comparative studies are so common, we often want to compare the proportions of two groups (such as men and women) that have some characteristic. We call the two groups being compared Population 1 and Population 2, and the two population proportions of “successes” p_1 and p_2 . The data consist of two independent SRSs, of size n_1 from Population 1 and size n_2 from Population 2. The proportion of successes in each sample estimates the corresponding population proportion. Here is the notation we will use in this section:

Population	Population proportion	Sample size	Count of successes	Sample proportion
1	p_1	n_1	X_1	$\hat{p}_1 = X_1/n_1$
2	p_2	n_2	X_2	$\hat{p}_2 = X_2/n_2$

To compare the two populations, we use the difference between the two sample proportions:

$$D = \hat{p}_1 - \hat{p}_2$$