CHAPTER 8 EXERCISES

8.54 "The nature of work is changing at whirlwind speed. Perhaps now more than ever before, job stress poses a threat to the health of workers and, in turn, to the health of organizations."[18] So says the National Institute for Occupational Safety and Health. Employers are concerned about the effect of stress on their employees. Stress can lower morale and efficiency and increase medical costs. A large survey of restaurant employees found that 75% reported that work stress had a negative impact on their personal lives.[19] The human resources manager of a chain of restaurants is concerned that work stress may be affecting the chain's employees. She asks a random sample of 100 employees to respond Yes or No to the question "Does work stress have a negative impact on your personal life?" Of these, 68 say "Yes." Give a 95% confidence interval for the proportion of employees who work for this chain of restaurants who believe that work stress has a negative impact on their personal lives.

8.55 Refer to the previous exercise. Is there evidence to conclude that the proportion for this chain of restaurants differs from the value given for the

national survey? For this exercise, assume that there is no error associ with the estimate for the national survey.

8.56 Following complaints about the working conditions in some apparel factories both in the United States and abroad, a joint government and industry commission recommended in 1998 that companies that moni and enforce proper standards be allowed to display a "No Sweat" label on their products. Does the presence of these labels influence consume behavior? A survey of U.S. residents aged 18 or older asked a series of questions about how likely they would be to purchase a garment unde various conditions. For some conditions, it was stated that the garmen a "No Sweat" label; for others, there was no mention of such a label. O the basis of the responses, each person was classified as a "label user" "label nonuser."[20] There were 296 women surveyed. Of these, 63 were users. On the other hand, 27 of 251 men were classified as users. Give confidence interval for the difference in the proportions.

8.57 Refer to the previous exercise. You would like to compare the women the men. Set up appropriate hypotheses, and find the test statistic and *P*-value. What do you conclude?

8.58 The Ping Company makes custom-built golf clubs and competes in the $4 billion golf equipment industry. To improve its business processes, P decided to seek ISO 9001 certification.[21] As part of this process, a study the time it took to repair golf clubs sent to the company by mail deter that 16% of orders were sent back to the customers in 5 days or less. P examined the processing of repair orders and made changes. Following changes, 90% of orders were completed within 5 days. Assume that ea the estimated percents is based on a random sample of 200 orders.

(a) How many orders were completed in 5 days or less before the cha Give a 95% confidence interval for the proportion of orders compl in this time.

(b) Do the same for orders after the changes.

(c) Give a 95% confidence interval for the improvement. Express this for a difference in proportions and for a difference in percents.

8.59 To devise effective marketing strategies it is helpful to know the characteristics of your customers. A study compared demographic characteristics of people who use the Internet for travel arrangements of people who do not.[22] Of 1132 Internet users, 643 had completed coll Among the 852 nonusers, 349 had completed college.

(a) Do users and nonusers differ significantly in the proportion of coll graduates?

(b) Give a 95% confidence interval for the difference in the proportion

8.60 The study mentioned in the previous exercise also asked about income. Among Internet users, 493 reported income of less than $50,000 and 37 reported income of $50,000 or more. (Not everyone answered the incon question.) The corresponding numbers for nonusers were 477 and 200.

Perform a significance test to compare the incomes of users with nonusers and also give an estimate of the difference in proportions with a 95% margin of error.

8.61 Refer to the previous two exercises. Give the total number of users and the total number of nonusers for the analysis of education. Do the same for the analysis of income. The difference is due to respondents who chose "Rather not say" for the income question. Give the proportions of "Rather not say" individuals for users and nonusers. Perform a significance test to compare these and give a 95% confidence interval for the difference. People are often reluctant to provide information about their income. Do you think that this type of nonresponse for the income is a serious limitation for this study?

8.62 According to literature on brand loyalty, consumers who are loyal to a brand are likely to consistently select the same product. This type of consistency could come from a positive childhood association. To examine brand loyalty among fans of the Chicago Cubs, 371 Cubs fans among patrons of a restaurant located in Wrigleyville were surveyed prior to a game at Wrigley Field, the Cubs home field.[23] The respondents were classified as "die-hard fans" or "less loyal fans." Of the 134 die-hard fans, 90.3% reported that they watched or listened to Cubs games when they were children. Among the 237 less loyal fans, 67.9% said that they watched or listened as children.

(a) Find the number of die-hard Cubs fans who watched or listened to games when they were children. Do the same for the less loyal fans.

(b) Use a significance test to compare the die-hard fans with the less loyal fans with respect to their childhood experiences relative to the team.

(c) Express the results with a 95% confidence interval for the difference in proportions.

8.63 The study mentioned in the previous exercise found that two-thirds of the die-hard fans attended Cubs games at least once a month, but only 20% of the less loyal fans attended this often. Analyze these data using a significance test and a confidence interval. Write a short summary of your findings.

8.64 A Gallup Poll used telephone interviews to survey a sample of 1025 U.S. residents over the age of 18 regarding their use of credit cards.[24] The poll reported that 76% of Americans said that they had at least one credit card. Give the 95% margin of error for this estimate.

8.65 The Gallup Poll in the previous exercise reported that 41% of those who have credit cards do not pay the full balance each month. Find the number of people in the survey who said that they had at least one credit card, using the information in the previous exercise. Combine this number with the reported 41% to give a margin of error for the proportion of credit card owners who do not pay their full balance.

8.66 Refer to the study of gender bias and stereotyping described in Exercise 8.35 (page 596). Here are the counts of "girl," "woman," "boy," and "man" for all

of the grammar texts studied. The one we analyzed in Exercise 8.35 was number 6.

	Text number									
	1	2	3	4	5	6	7	8	9	10
Girl	2	5	25	11	2	48	38	5	48	13
Woman	3	2	31	65	1	12	2	13	24	5
Boy	7	18	14	19	12	52	70	6	128	32
Man	27	45	51	138	31	80	2	27	48	95

For each text perform the significance test to compare the proportions of juvenile references for females and males. Summarize the results of the significance tests for the 10 texts studied. The researchers who conducted the study note that the authors of the last three texts are women, while the other seven texts were written by men. Do you see any pattern that suggests that the gender of the author is associated with the results?

8.67 Refer to the previous exercise. Let us now combine the categories "girl" "woman" and "boy" with "man." For each text calculate the proportion of male references and test the hypothesis that male and female references are equally likely (that is, the proportion of male references is equal to $\frac{1}{2}$). Summarize the results of your 10 tests. Is there a pattern that suggests a relation with the gender of the author?

8.68 Many new products introduced into the market are targeted toward children. The choice behavior of children with regard to new products is of particular interest to companies that design marketing strategies for these products. As part of one study, children in different age groups were compared on their ability to sort new products into the correct product category (milk or juice).[25] Here are some of the data:

Age group	n	Number who sorted correctly
4- to 5-year-olds	50	10
6- to 7-year-olds	53	28

Test the null hypothesis that the two age groups are equally skilled at sorting. Justify your choice of an alternative hypothesis. Also, give a 90% confidence interval for the difference. Summarize your results with a short paragraph.

8.69 A television news program conducts a call-in poll about a proposed city ordinance on handgun ownership. Of the 2372 calls, 1921 oppose the ban. The station, following recommended practice, makes a confidence statement: "81% of the Channel 13 Pulse Poll sample opposed the ban. We can be 95%

confident that the true proportion of citizens opposing a handgun ban is within 1.6% of the sample result." Is this conclusion justified?

8.70 Eleven percent of the products produced by an industrial process over the past several months fail to conform to the specifications. The company modifies the process in an attempt to reduce the rate of nonconformities. In a trial run, the modified process produces 16 nonconforming items out of a total of 300 produced. Do these results demonstrate that the modification is effective? Support your conclusion with a clear statement of your assumptions and the results of your statistical calculations.

8.71 In the setting of the previous exercise, give a 95% confidence interval for the proportion of nonconforming items for the modified process. Then, taking $p_0 = 0.11$ to be the old proportion and p the proportion for the modified process, give a 95% confidence interval for $p - p_0$.

8.72 In a random sample of 950 students from a large public university, it was found that 444 of the students changed majors during their college years.

(a) Give a 95% confidence interval for the proportion of students at this university who change majors.

(b) Express your results from (a) in terms of the *percent* of students who change majors.

(c) University officials concerned with counseling students are interested in the number of students who change majors rather than the proportion. The university has 38,000 undergraduate students. Convert the confidence interval you found in (a) to a confidence interval for the *number* of students who change majors during their college years.

8.73 In a study on blood pressure and diet, a random sample of Seventh-Day Adventists were interviewed at a national meeting. Because many people who belong to this denomination are vegetarians, they are a very useful group for studying the effects of a meatless diet.[26] Blacks in the population as a whole have a higher average blood pressure than whites. A study of this type should therefore take race into account in the analysis. The 312 people in the sample were categorized by race and whether or not they were vegetarians. The data are given in the following table:

	Black	White
Vegetarian	42	135
Not vegetarian	47	88

Are the proportions of vegetarians the same among all black and white Seventh-Day Adventists who attended this meeting? Analyze the data, paying particular attention to this question. Summarize your analysis and conclusions. What can you infer about the proportions of vegetarians among black and white Seventh-Day Adventists in general? What about blacks and whites in general?

8.74 In Example 8.11 (page 594) we discussed a study that examined the association between high blood pressure and increased risk of death f cardiovascular disease. There were 2676 men with low blood pressure 3338 men with high blood pressure. In the low-blood-pressure group, men died from cardiovascular disease; in the high-blood-pressure gro 55 died.

(a) Verify the calculations in Example 8.11 by computing the 95% confidence interval for the difference in proportions.

(b) Do the study data confirm that death rates are higher among me high blood pressure? State hypotheses, carry out a significance te and give your conclusions.

8.75 An experiment designed to assess the effects of aspirin on cardiova disease studied 5139 male British medical doctors. The doctors w randomly assigned to two groups. One group of 3429 doctors took aspirin daily, and the other group did not take aspirin. After 6 years were 148 deaths from heart attack or stroke in the first group and 7 the second group. A similar experiment used male American medi doctors as subjects. These doctors were also randomly assigned to two groups. The 11,037 doctors in the first group took one aspirin other day, and the 11,034 doctors in the second group took no asp After nearly 5 years, there were 104 deaths from heart attacks in th group and 189 in the second.[27] Analyze the data from these two studi summarize the results. How do the conclusions of the two studies and why?

8.76 Gastric freezing was once a recommended treatment for ulcers in the intestine. A randomized comparative experiment found that 28 of the patients who were subjected to gastric freezing improved, while 30 of patients in the control group improved.

(a) State the appropriate hypothesis and a two-sided alternative and out a z test. What is the P-value?

(b) What do you conclude about the effectiveness of gastric freezing treatment for ulcers? (See Example 3.5 on page 231 for a discussi gastric freezing.)

8.77 In this exercise we examine the effect of the sample size on the signifi test for comparing two proportions. In each case suppose that $\hat{p}_1 = 0.$ and $\hat{p}_2 = 0.5$, and take n to be the common value of n_1 and n_2. Use th statistic to test H_0: $p_1 = p_2$ versus the alternative H_a: $p_1 \neq p_2$. Comput statistic and the associated P-value for the following values of n: 10, 2 60, 80, 100, and 500. Summarize the results in a table. Explain what y observe about the effect of the sample size on statistical significance w the sample proportions $\hat{p}_1$ and $\hat{p}_2$ are unchanged.

8.78 In the first section of this chapter, we studied the effect of the sample s on the margin of error of the confidence interval for a single proportio this exercise we perform some calculations to observe this effect for th

two-sample problem. Suppose that $\tilde{p}_1 = 0.7$ and $\tilde{p}_2 = 0.5$, and n represents the common value of n_1 and n_2. Compute the 95% margins of error for the difference in the two proportions for $n = 16$, 26, 56, 86, 106, and 506. Summarize and explain your results.

8.79 For a single proportion the margin of error of a confidence interval is largest for any given sample size n and confidence level C when $\hat{p} = 0.5$. This led us to use $p^* = 0.5$ for planning purposes. The same kind of result is true for the two-sample problem. The margin of error of the confidence interval for the difference between two proportions is largest when $\hat{p}_1 = \hat{p}_2 = 0.5$. Use these conservative values in the following calculations, and assume that the sample sizes n_1 and n_2 have the common value n. Calculate the margins of error of the 95% confidence intervals for the difference in two proportions for the following choices of n: 16, 26, 56, 86, 106, and 506. Present the results in a table or with a graph. Summarize your conclusions.

8.80 As the previous problem noted, using the guessed value 0.5 for both $\tilde{p}_1$ and $\tilde{p}_2$ gives a conservative margin of error in confidence intervals for the difference between two population proportions. You are planning a survey and will calculate a 95% confidence interval for the difference in two proportions when the data are collected. You would like the margin of error of the interval to be less than or equal to 0.05. You will use the same sample size n for both populations.

(a) How large a value of n is needed?

(b) Give a general formula for n in terms of the desired margin of error m and the critical value z^*.

8.81 *Castaneda v. Partida* is an important court case in which statistical methods were used as part of a legal argument.[28] When reviewing this case, the Supreme Court used the phrase "two or three standard deviations" as a criterion for statistical significance. This Supreme Court review has served as the basis for many subsequent applications of statistical methods in legal settings. (The two or three standard deviations referred to by the Court are values of the z statistic and correspond to P-values of approximately 0.05 and 0.0026.) In *Castaneda* the plaintiffs alleged that the method for selecting juries in a county in Texas was biased against Mexican Americans. For the period of time at issue, there were 181,535 persons eligible for jury duty, of whom 143,611 were Mexican Americans. Of the 870 people selected for jury duty, 339 were Mexican Americans.

(a) What proportion of eligible voters were Mexican Americans? Let this value be p_0.

(b) Let p be the probability that a randomly selected juror is a Mexican American. The null hypothesis to be tested is H_0: $p = p_0$. Find the value of $\hat{p}$ for this problem, compute the z statistic, and find the P-value. What do you conclude? (A finding of statistical significance in this circumstance does not constitute a proof of discrimination. It can

be used, however, to establish a prima facie case. The burden of pr then shifts to the defense.)

(c) We can reformulate this exercise as a two-sample problem. Here w wish to compare the proportion of Mexican Americans among tho selected as jurors with the proportion of Mexican Americans amor those not selected as jurors. Let p_1 be the probability that a rando selected juror is a Mexican American, and let p_2 be the probabilit that a randomly selected nonjuror is a Mexican American. Find th z statistic and its P-value. How do your answers compare with you results in (b)?