- Confidence intervals provide additional information that significance tests do not—namely, a set of plausible values for the true population parameter p. A two-sided test of $H_0: p = p_0$ at significance level α gives roughly the same conclusion as a $100(1 - \alpha)\%$ confidence interval.
- The **power** of a significance test against a specific alternative is the probability that the test will reject H_0 when the alternative is true. Power measures the ability of the test to detect an alternative value of the parameter. For a specific alternative, $P(\text{Type II error}) = 1 - \text{power}$.
- There is an important link between the probabilities of Type I and Type II error in a significance test: as one increases, the other decreases.
- We can increase the power of a significance test by increasing the sample size, increasing the significance level, or increasing the difference that is important to detect between the null and alternative parameter values.

9.2 TECHNOLOGY CORNER

TI-Nspire Instructions in Appendix B; HP Prime instructions on the book's Web site.

18. One-proportion z test on the calculator page 561

Section 9.2 Exercises

In Exercises 31 and 32, check that the conditions for carrying out a one-sample z test for the population proportion p are met.

31. **Home computers** Jason reads a report that says 80% of U.S. high school students have a computer at home. He believes the proportion is smaller than 0.80 at his large rural high school. Jason chooses an SRS of 60 students and records whether they have a computer at home. pg 555

32. **Walking to school** A recent report claimed that 13% of students typically walk to school.[10] DeAnna thinks that the proportion is higher than 0.13 at her large elementary school, so she surveys a random sample of 100 students to find out.

In Exercises 33 and 34, explain why we aren't safe carrying out a one-sample z test for the population proportion p.

33. **No test** You toss a coin 10 times to perform a test of $H_0: p = 0.5$ that the coin is balanced against $H_a: p \neq 0.5$.

34. **No test** A college president says, "99% of the alumni support my firing of Coach Boggs." You contact an SRS of 200 of the college's 15,000 living alumni to perform a test of $H_0: p = 0.99$ versus $H_a: p < 0.99$.

35. **Home computers** Refer to Exercise 31. In Jason's SRS, 41 of the students had a computer at home. pg 556
    (a) Calculate the test statistic.
    (b) Find the P-value using Table A or technology. Show this result as an area under a standard Normal curve.

36. **Walking to school** Refer to Exercise 32. For DeAnna's survey, 17 students in the sample said they typically walk to school.
    (a) Calculate the test statistic.
    (b) Find the P-value using Table A or technology. Show this result as an area under a standard Normal curve.

37. **Significance tests** A test of $H_0: p = 0.5$ versus $H_a: p > 0.5$ has test statistic $z = 2.19$.
    (a) What conclusion would you draw at the 5% significance level? At the 1% level?
    (b) If the alternative hypothesis were $H_a: p \neq 0.5$, what conclusion would you draw at the 5% significance level? At the 1% level?

38. **Significance tests** A test of $H_0: p = 0.65$ against $H_a: p < 0.65$ has test statistic $z = -1.78$.

(a) What conclusion would you draw at the 5% significance level? At the 1% level?

(b) If the alternative hypothesis were $H_a: p \neq 0.65$, what conclusion would you draw at the 5% significance level? At the 1% level?

39. **Better parking** A local high school makes a change that should improve student satisfaction with the parking situation. Before the change, 37% of the school's students approved of the parking that was provided. After the change, the principal surveys an SRS of 200 of the over 2500 students at the school. In all, 83 students say that they approve of the new parking arrangement. The principal cites this as evidence that the change was effective. Perform a test of the principal's claim at the $\alpha = 0.05$ significance level.

pg 559

40. **Side effects** A drug manufacturer claims that less than 10% of patients who take its new drug for treating Alzheimer's disease will experience nausea. To test this claim, researchers conduct an experiment. They give the new drug to a random sample of 300 out of 5000 Alzheimer's patients whose families have given informed consent for the patients to participate in the study. In all, 25 of the subjects experience nausea. Use these data to perform a test of the drug manufacturer's claim at the $\alpha = 0.05$ significance level.

41. **Are boys more likely?** We hear that newborn babies are more likely to be boys than girls. Is this true? A random sample of 25,468 firstborn children included 13,173 boys.[11]

(a) Do these data give convincing evidence that firstborn children are more likely to be boys than girls?

(b) To what population can the results of this study be generalized: all children or all firstborn children? Justify your answer.

42. **Fresh coffee** People of taste are supposed to prefer fresh-brewed coffee to the instant variety. On the other hand, perhaps many coffee drinkers just want their caffeine fix. A skeptic claims that only half of all coffee drinkers prefer fresh-brewed coffee. To test this claim, we ask a random sample of 50 coffee drinkers in a small city to take part in a study. Each person tastes two unmarked cups—one containing instant coffee and one containing fresh-brewed coffee—and says which he or she prefers. We find that 36 of the 50 choose the fresh coffee.

(a) Do these results give convincing evidence that coffee drinkers favor fresh-brewed over instant coffee?

(b) We presented the two cups to each coffee drinker in a random order, so that some people tasted the fresh coffee first, while others drank the instant coffee first. Why do you think we did this?

43. **Bullies in middle school** A University of Illinois study on aggressive behavior surveyed a random sample of 558 middle school students. When asked to describe their behavior in the last 30 days, 445 students said their behavior included physical aggression, social ridicule, teasing, name-calling, and issuing threats. This behavior was not defined as bullying in the questionnaire.[12] Is this evidence that more than three-quarters of middle school students engage in bullying behavior? To find out, Maurice decides to perform a significance test. Unfortunately, he made a few errors along the way. Your job is to spot the mistakes and correct them.

$$H_0: p = 0.75$$
$$H_a: \hat{p} > 0.797$$

where p = the true mean proportion of middle school students who engaged in bullying.

- A random sample of 558 middle school students was surveyed.
- $558(0.797) = 444.73$ is at least 10.

$$z = \frac{0.75 - 0.797}{\frac{\sqrt{0.797(0.203)}}{445}} = -2.46; \text{ } P\text{-value} = 2(0.0069) = 0.0138$$

The probability that the null hypothesis is true is only 0.0138, so we reject H_0. This proves that more than three-quarters of the school engaged in bullying behavior.

44. **Is this coin fair?** The French naturalist Count Buffon (1707–1788) tossed a coin 4040 times. He got 2048 heads. That's a bit more than one-half. Is this evidence that Count Buffon's coin was not balanced? To find out, Luisa decides to perform a significance test. Unfortunately, she made a few errors along the way. Your job is to spot the mistakes and correct them.

$$H_0: \mu > 0.5$$
$$H_a: \bar{x} = 0.5$$

- 10%: $4040(0.5) = 2020$ and $4040(1 - 0.5) = 2020$ are both at least 10.
- Large Counts: There are at least 40,400 coins in the world.

$$t = \frac{0.5 - 0.507}{\sqrt{\frac{0.5(0.5)}{4040}}} = -0.89; \text{ } P\text{-value} = 1 - 0.1867 = 0.8133$$

Reject H_0 because the P-value is so large and conclude that the coin is fair.

45. **Teen drivers** A state's Division of Motor Vehicles (DMV) claims that 60% of teens pass their driving test on the first attempt. An investigative reporter examines an SRS of the DMV records for 125 teens; 86 of them passed the test on their first try. Is there convincing evidence at the $\alpha = 0.05$ significance level that the DMV's claim is incorrect?

pg 562

46. **We want to be rich** In a recent year, 73% of first-year college students responding to a national survey identified "being very well-off financially" as an important personal goal. A state university finds that 132 of an SRS of 200 of its first-year students say that this goal is important. Is there convincing evidence at the $\alpha = 0.05$ significance level that the proportion of all first-year students at this university who think being very well-off is important differs from the national value, 73%?

47. **Teen drivers** Refer to Exercise 45.

(a) Construct and interpret a 95% confidence interval for the proportion of all teens in the state who passed their driving test on the first attempt.

(b) Explain what the interval in part (a) tells you about the DMV's claim.

48. **We want to be rich** Refer to Exercise 46.

(a) Construct and interpret a 95% confidence interval for the true proportion p of all first-year students at the university who would identify being well-off as an important personal goal.

(b) Explain what the interval in part (a) tells you about whether the national value holds at this university.

49. **Do you Tweet?** In early 2012, the Pew Internet and American Life Project asked a random sample of U.S. adults, "Do you ever . . . use Twitter or another service to share updates about yourself or to see updates about others?" According to Pew, the resulting 95% confidence interval is (0.123, 0.177).[13] Does this interval provide convincing evidence that the actual proportion of U.S. adults who would say they use Twitter differs from 0.16? Justify your answer.

50. **Losing weight** A Gallup Poll found that 59% of the people in its sample said "Yes" when asked, "Would you like to lose weight?" Gallup announced: "For results based on the total sample of national adults, one can say with 95% confidence that the margin of (sampling) error is ±3 percentage points."[14] Does this interval provide convincing evidence that the actual proportion of U.S. adults who would say they want to lose weight differs from 0.55? Justify your answer.

51. **Teens and sex** The Gallup Youth Survey asked a random sample of U.S. teens aged 13 to 17 whether they thought that young people should wait to have sex until marriage.[15] The Minitab output below shows the results of a significance test and a 95% confidence interval based on the survey data.

```
Session
Test and CI for One Proportion
Test of p = 0.5 vs p not = 0.5

Sample    X    N  Sample p           95% CI            Z-Value  P-Value
1       246  439  0.560364  (0.513935, 0.606794)      2.53    0.011
```

(a) Define the parameter of interest.

(b) Check that the conditions for performing the significance test are met in this case.

(c) Interpret the P-value in context.

(d) Do these data give convincing evidence that the actual population proportion differs from 0.5? Justify your answer with appropriate evidence.

52. **Reporting cheating** What proportion of students are willing 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?" The Minitab output below shows the results of a significance test and a 95% confidence interval based on the survey data.[16]

```
Session
Test and CI for One Proportion
Test of p = 0.15 vs p not = 0.15

Sample   X    N  Sample p           95% CI            Z-Value  P-Value
1       19  172  0.110465  (0.063619, 0.157312)     -1.45    0.146
```

(a) Define the parameter of interest.

(b) Check that the conditions for performing the significance test are met in this case.

(c) Interpret the P-value in context.

(d) Do these data give convincing evidence that the actual population proportion differs from 0.15? Justify your answer with appropriate evidence.

53. **Better parking** Refer to Exercise 39.

(a) Describe a Type I error and a Type II error in this setting, and explain the consequences of each.

(b) The test has a power of 0.75 to detect that $p = 0.45$. Explain what this means.

(c) Identify two ways to increase the power in part (b).

54. **Side effects** Refer to Exercise 40.

(a) Describe a Type I error and a Type II error in this setting, and explain the consequences of each.

(b) The test has a power of 0.54 to detect that $p = 0.07$. Explain what this means.

(c) Identify two ways to increase the power in part (b).

55. **Error probabilities** You read that a statistical test at significance level $\alpha = 0.05$ has power 0.78. What are the probabilities of Type I and Type II errors for this test?

56. **Error probabilities** You read that a statistical test at the $\alpha = 0.01$ level has probability 0.14 of making a Type II error when a specific alternative is true. What is the power of the test against this alternative?

57. **Power** A drug manufacturer claims that fewer than 10% of patients who take its new drug for treating Alzheimer's disease will experience nausea. To test this claim, a significance test is carried out of

$$H_0: p = 0.10$$
$$H_a: p < 0.10$$

You learn that the power of this test at the 5% significance level against the alternative $p = 0.08$ is 0.29.

(a) Explain in simple language what "power = 0.29" means in this setting.

(b) You could get higher power against the same alternative with the same α by changing the number of measurements you make. Should you make more measurements or fewer to increase power? Explain.

(c) If you decide to use $\alpha = 0.01$ in place of $\alpha = 0.05$, with no other changes in the test, will the power increase or decrease? Justify your answer.

(d) If you shift your interest to the alternative $p = 0.07$ with no other changes, will the power increase or decrease? Justify your answer.

58. **What is power?** You manufacture and sell a liquid product whose electrical conductivity is supposed to be 5. You plan to make six measurements of the conductivity of each lot of product. If the product meets specifications, the mean of many measurements will be 5. You will therefore test

$$H_0: \mu = 5$$
$$H_a: \mu \neq 5$$

If the true conductivity is 5.1, the liquid is not suitable for its intended use. You learn that the power of your test at the 5% significance level against the alternative $\mu = 5.1$ is 0.23.

(a) Explain in simple language what "power = 0.23" means in this setting.

(b) You could get higher power against the same alternative with the same α by changing the number of measurements you make. Should you make more measurements or fewer to increase power?

(c) If you decide to use $\alpha = 0.10$ in place of $\alpha = 0.05$, with no other changes in the test, will the power increase or decrease? Justify your answer.

(d) If you shift your interest to the alternative $\mu = 5.2$, with no other changes, will the power increase or decrease? Justify your answer.

Multiple choice: Select the best answer for Exercises 59 to 62.

59. After once again losing a football game to the archrival, a college's alumni association conducted a survey to see if alumni were in favor of firing the coach. An SRS of 100 alumni from the population of all living alumni was taken, and 64 of the alumni in the sample were in favor of firing the coach. Suppose you wish to see if a majority of living alumni are in favor of firing the coach. The appropriate test statistic is

(a) $z = \dfrac{0.64 - 0.5}{\sqrt{\dfrac{0.64(0.36)}{100}}}$

(b) $t = \dfrac{0.64 - 0.5}{\sqrt{\dfrac{0.64(0.36)}{100}}}$

(c) $z = \dfrac{0.64 - 0.5}{\sqrt{\dfrac{0.5(0.5)}{100}}}$

(d) $z = \dfrac{0.64 - 0.5}{\sqrt{\dfrac{0.64(0.36)}{64}}}$

(e) $z = \dfrac{0.5 - 0.64}{\sqrt{\dfrac{0.5(0.5)}{100}}}$

60. Which of the following is *not* a condition for performing a significance test about a population proportion p?

(a) The data should come from a random sample or randomized experiment.

(b) Both np_0 and $n(1 - p_0)$ should be at least 10.

(c) If you are sampling without replacement from a finite population, then you should sample no more than 10% of the population.

(d) The population distribution should be approximately Normal, unless the sample size is large.

(e) All of the above are conditions for performing a significance test about a population proportion.

61. The z statistic for a test of $H_0: p = 0.4$ versus $H_a: p \neq 0.4$ is $z = 2.43$. This test is

(a) not significant at either $\alpha = 0.05$ or $\alpha = 0.01$.

(b) significant at $\alpha = 0.05$ but not at $\alpha = 0.01$.

(c) significant at $\alpha = 0.01$ but not at $\alpha = 0.05$.

(d) significant at both $\alpha = 0.05$ and $\alpha = 0.01$.

(e) inconclusive because we don't know the value of $\hat{p}$.

62. Which of the following 95% confidence intervals would lead us to reject $H_0: p = 0.30$ in favor of $H_a: p \neq 0.30$ at the 5% significance level?

(a) (0.19, 0.27) (c) (0.27, 0.31) (e) None of these

(b) (0.24, 0.30) (d) (0.29, 0.38)

63. **Packaging CDs** (6.2, 5.3) A manufacturer of compact discs (CDs) wants to be sure that their CDs will fit inside the plastic cases they have bought for packaging. Both the CDs and the cases are circular. According to the supplier, the plastic cases vary Normally with mean diameter $\mu = 4.2$ inches and standard deviation $\sigma = 0.05$ inches. The CD manufacturer decides to produce CDs with mean diameter $\mu = 4$ inches. Their diameters follow a Normal distribution with $\sigma = 0.1$ inches.

(a) Let X = the diameter of a randomly selected CD and Y = the diameter of a randomly selected case. Describe the shape, center, and spread of the distribution of the random variable $X - Y$. What is the importance of this random variable to the CD manufacturer?

(b) Compute the probability that a randomly selected CD will fit inside a randomly selected case.

(c) The production process actually runs in batches of 100 CDs. If each of these CDs is paired with a randomly chosen plastic case, find the probability that all the CDs fit in their cases.

64. **Cash to find work?** (4.2) Will cash bonuses speed the return to work of unemployed people? The Illinois Department of Employment Security designed an experiment to find out. The subjects were 10,065 people aged 20 to 54 who were filing claims for unemployment insurance. Some were offered \$500 if they found a job within 11 weeks and held it for at least 4 months. Others could tell potential employers that the state would pay the employer \$500 for hiring them. A control group got neither kind of bonus.[17]

(a) Describe a completely randomized design for this experiment.

(b) How will you label the subjects for random assignment? Use Table D at line 127 to choose the first 3 subjects for the first treatment.

(c) Explain the purpose of a control group in this setting.

9.3 Tests about a Population Mean

...T YOU WILL LEARN By the end of the section, you should be able to:

- ...heck the Random, 10%, and Normal/Large ... for performing a significance test
- ...t a population mean.
- Use a confidence interval to draw a conclusion for a two-sided test about a population parameter.
- Perform a significance test about a mean difference using paired data.

...e intervals and significance tests for a population proportion p are ... z-values from the standard Normal distribution. Inference about ...ation mean μ uses a t distribution with $n - 1$ degrees of freedom, ... in the rare case when the population standard deviation σ is known. ...earned how to construct confidence intervals for a population mean in ...tion 8.3. Now we'll examine the details of testing a claim about a population mean μ.