10.5 Let Y_1 and Y_2 be independent and identically distributed with a uniform distribution over the interval $(\theta, \theta + 1)$. For testing $H_0 : \theta = 0$ versus $H_a : \theta > 0$, we have two competing tests:

Test 1: Reject H_0 if $Y_1 > .95$.
Test 2: Reject H_0 if $Y_1 + Y_2 > c$.

Find the value of c so that test 2 has the same value for α as test 1. [*Hint*: In Example 6.3, we derived the density and distribution function of the sum of two independent random variables that are uniformly distributed on the interval (0, 1).]

10.6 We are interested in testing whether or not a coin is balanced based on the number of heads Y on 36 tosses of the coin. ($H_0 : p = .5$ versus $H_a : p \neq .5$). If we use the rejection region $|y - 18| \geq 4$, what is

a the value of α?
b the value of β if $p = .7$?

10.7 **True or False** Refer to Exercise 10.6.

a The level of the test computed in Exercise 10.6(a) is the probability that H_0 is true.
b The value of β computed in Exercise 10.6(b) is the probability that H_a is true.
c In Exercise 10.6(b), β was computed assuming that the null hypothesis was false.
d If β was computed when $p = 0.55$, the value would be larger than the value of β obtained in Exercise 10.6(b).
e The probability that the test mistakenly rejects H_0 is β.
f Suppose that RR was changed to $|y - 18| \geq 2$.
  i This RR would lead to rejecting the null hypothesis more often than the RR used in Exercise 10.6.
  ii If α was computed using this new RR, the value would be larger than the value obtained in Exercise 10.6(a).
  iii If β was computed when $p = .7$ and using this new RR, the value would be larger than the value obtained in Exercise 10.6(b).

***10.8** A two-stage clinical trial is planned for testing $H_0 : p = .10$ versus $H_a : p > .10$, where p is the proportion of responders among patients who were treated by the protocol treatment. At the first stage, 15 patients are accrued and treated. If 4 or more responders are observed among the (first) 15 patients, H_0 is rejected, the study is terminated, and no more patients are accrued. Otherwise, another 15 patients will be accrued and treated in the second stage. If a total of 6 or more responders are observed among the 30 patients accrued in the two stages (15 in the first stage and 15 more in the second stage), then H_0 is rejected. For example, if 5 responders are found among the first-stage patients, H_0 is rejected and the study is over. However, if 2 responders are found among the first-stage patients, 15 second-stage patients are accrued, and an additional 4 or more responders (for a total of 6 or more among the 30) are identified, H_0 is rejected and the study is over.[1]

a Use the binomial table to find the numerical value of α for this testing procedure.
b Use the binomial table to find the probability of rejecting the null hypothesis when using this rejection region if $p = .30$.
c For the rejection region defined above, find β if $p = .30$.

1. Exercises preceded by an asterisk are optional.

c Change the real value of p to .2 and simulate at least 200 tests. Click the button "Show Summary." Does anything look wrong?

10.17 A survey published in the *American Journal of Sports Medicine*[2] reported the number of meters (m) per week swum by two groups of swimmers—those who competed exclusively in breaststroke and those who competed in the individual medley (which includes breaststroke). The number of meters per week practicing the breaststroke was recorded for each swimmer, and the summary statistics are given below. Is there sufficient evidence to indicate that the average number of meters per week spent practicing breaststroke is greater for exclusive breaststrokers than it is for those swimming individual medley?

	Specialty	
	Exclusively Breaststroke	Individual Medley
Sample size	130	80
Sample mean (m)	9017	5853
Sample standard deviation (m)	7162	1961
Population mean	μ_1	μ_2

a State the null and alternative hypotheses.

b What is the appropriate rejection region for an $\alpha = .01$ level test?

c Calculate the observed value of the appropriate test statistic.

d What is your conclusion?

e What is a practical reason for the conclusion you reached in part (d)?

10.18 The hourly wages in a particular industry are normally distributed with mean \$13.20 and standard deviation \$2.50. A company in this industry employs 40 workers, paying them an average of \$12.20 per hour. Can this company be accused of paying substandard wages? Use an $\alpha = .01$ level test.

10.19 The output voltage for an electric circuit is specified to be 130. A sample of 40 independent readings on the voltage for this circuit gave a sample mean 128.6 and standard deviation 2.1. Test the hypothesis that the average output voltage is 130 against the alternative that it is less than 130. Use a test with level .05.

10.20 The Rockwell hardness index for steel is determined by pressing a diamond point into the steel and measuring the depth of penetration. For 50 specimens of an alloy of steel, the Rockwell hardness index averaged 62 with standard deviation 8. The manufacturer claims that this alloy has an average hardness index of at least 64. Is there sufficient evidence to refute the manufacturer's claim at the 1% significance level?

10.21 Shear strength measurements derived from unconfined compression tests for two types of soils gave the results shown in the following table (measurements in tons per square foot). Do the soils appear to differ with respect to average shear strength, at the 1% significance level?

Soil Type I	Soil Type II
$n_1 = 30$	$n_2 = 35$
$\overline{y}_1 = 1.65$	$\overline{y}_2 = 1.43$
$s_1 = 0.26$	$s_2 = 0.22$

2. *Source:* Kurt Grote, T. L. Lincoln, and J. G. Gamble, "Hip Adductor Injury in Competitive Swimmers," *American Journal of Sports Medicine* 32(1) (2004): 104.

10.22 In Exercise 8.66, we examined the results of a 2001 study by Leonard, Speziale and Pernick comparing traditional and activity-oriented methods for teaching biology. Pretests were given to students who were subsequently taught by one of the two methods. Summary statistics were given for the pretest scores for 368 students who were subsequently taught using the traditional method and 372 who were taught using the activity-oriented method.

a Without looking at the data, would you expect there to be a difference in the mean *pretest* scores for those subsequently taught using the different methods? Based on your conjecture, what alternative hypothesis would you choose to test versus the null hypothesis that there is no difference in the mean pretest scores for the two groups?

b Does the alternative hypothesis that you posed in part (a) correspond to a one-tailed or a two-tailed statistical test?

c The mean and standard deviation of the pretest scores for those subsequently taught using the traditional method were 14.06 and 5.45, respectively. For those subsequently taught using the activity-oriented method, the respective corresponding mean and standard deviation were 13.38 and 5.59. Do the data provide support for the conjecture that the mean pretest scores do not differ for students subsequently taught using the two methods? Test using $\alpha = .01$.

10.23 Studies of the habits of white-tailed deer indicate that these deer live and feed within very limited ranges, approximately 150 to 205 acres. To determine whether the ranges of deer located in two different geographical areas differ, researchers caught, tagged, and fitted 40 deer with small radio transmitters. Several months later, the deer were tracked and identified, and the distance y from the release point was recorded. The mean and standard deviation of the distances from the release point were as given in the accompanying table.[3]

	Location	
	1	2
Sample size	40	40
Sample mean (ft)	2980	3205
Sample standard deviation (ft)	1140	963
Population mean	μ_1	μ_2

a If you have no preconceived reason for believing that one population mean is larger than the other, what would you choose for your alternative hypothesis? Your null hypothesis?

b Would your alternative hypothesis in part (a) imply a one-tailed or a two-tailed test? Explain.

c Do the data provide sufficient evidence to indicate that the mean distances differ for the two geographical locations? Test using $\alpha = .10$.

10.24 A study by Children's Hospital in Boston indicates that about 67% of American adults and about 15% of children and adolescents are overweight.[4] Thirteen children in a random sample of size 100 were found to be overweight. Is there sufficient evidence to indicate that the percentage reported by Children's Hospital is too high? Test at the $\alpha = 0.05$ level of significance.

10.25 An article in *American Demographics* reports that 67% of American adults always vote in presidential elections.[5] To test this claim, a random sample of 300 adults was taken, and 192

3. *Source:* Charles Dickey, "A Strategy for Big Bucks," *Field and Stream*, October 1990.

4. *Source:* Judy Holland, "'Cheeseburger Bill' on the Menu," *Press-Enterprise* (Riverside, Calif.), March 9, 2004, p. E1.

5. *Source:* Christopher Reynolds, "Rocking the Vote," *American Demographics*, February 2004, p. 48.

stated that they always voted in presidential elections. Do the results of this sample provide sufficient evident to indicate that the percentage of adults who say that they always vote in presidential elections is different than the percentage reported in *American Demographics*? Test using $\alpha = .01$.

10.26 According to the *Washington Post*, nearly 45% of all Americans are born with brown eyes, although their eyes don't necessarily stay brown.[6] A random sample of 80 adults found 32 with brown eyes. Is there sufficient evidence at the .01 level to indicate that the proportion of brown-eyed adults differs from the proportion of Americans who are born with brown eyes?

10.27 The state of California is working very hard to ensure that all elementary age students whose native language is not English become proficient in English by the sixth grade. Their progress is monitored each year using the California English Language Development test. The results for two school districts in southern California for the 2003 school year are given in the accompanying table.[7] Do the data indicate a significant difference in the 2003 proportions of students who are fluent in English for the two districts? Use $\alpha = .01$.

District	Riverside	Palm Springs
Number of students tested	6124	5512
Percentage fluent	40	37

10.28 The commercialism of the U.S. space program has been a topic of great interest since Dennis Tito paid \$20 million to ride along with the Russian cosmonauts on the space shuttle.[8] In a survey of 500 men and 500 women, 20% of the men and 26% of the women responded that space should remain commercial free.

a Does statistically significant evidence exist to suggest that there is a difference in the population proportions of men and women who think that space should remain commercial free? Use a .05 level test.

b Why is a statistically significant difference in these population proportions of *practical* importance to advertisers?

10.29 A manufacturer of automatic washers offers a model in one of three colors: A, B, or C. Of the first 1000 washers sold, 400 were of color A. Would you conclude that customers have a preference for color A? Justify your answer.

10.30 A manufacturer claimed that at least 20% of the public preferred her product. A sample of 100 persons is taken to check her claim. With $\alpha = .05$, how small would the sample percentage need to be before the claim could legitimately be refuted? (Notice that this would involve a one-tailed test of the hypothesis.)

10.31 What conditions must be met for the Z test to be used to test a hypothesis concerning a population mean μ?

10.32 In March 2001, a Gallup poll asked, "How would you rate the overall quality of the environment in this country today—as excellent, good, fair or poor?" Of 1060 adults nationwide, 46% gave a rating of excellent or good. Is this convincing evidence that a majority of the nation's adults think the quality of the environment is fair or poor? Test using $\alpha = .05$.

6. *Source:* "Seeing the World Through Tinted Lenses," *Washington Post*, March 16, 1993, p. 5.

7. *Source:* Cadonna Peyton, "Pupils Build English Skills," *Press-Enterprise* (Riverside, Calif.), March 19, 2004, p. B-1.

8. *Source:* Adapted from "Toplines: To the Moon?" *American Demographics*, August 2001, p. 9.

responsibility of choosing α and, possibly, the problem of evaluating the probability β of making a type II error are shifted to the reader.

Exercises

10.50 High airline occupancy rates on scheduled flights are essential for profitability. Suppose that a scheduled flight must average at least 60% occupancy to be profitable and that an examination of the occupancy rates for 120 10:00 A.M. flights from Atlanta to Dallas showed mean occupancy rate per flight of 58% and standard deviation 11%. Test to see if sufficient evidence exists to support a claim that the flight is unprofitable. Find the p-value associated with the test. What would you conclude if you wished to implement the test at the $\alpha = .10$ level?

10.51 Two sets of elementary schoolchildren were taught to read by using different methods, 50 by each method. At the conclusion of the instructional period, a reading test yielded the results $\overline{y}_1 = 74$, $\overline{y}_2 = 71$, $s_1 = 9$, and $s_2 = 10$.

a What is the attained significance level if you wish to see whether evidence indicates a difference between the two population means?

b What would you conclude if you desired an α-value of .05?

10.52 A biologist has hypothesized that high concentrations of actinomycin D inhibit RNA synthesis in cells and thereby inhibit the production of proteins. An experiment conducted to test this theory compared the RNA synthesis in cells treated with two concentrations of actinomycin D: 0.6 and 0.7 micrograms per liter. Cells treated with the lower concentration (0.6) of actinomycin D yielded that 55 out of 70 developed normally whereas only 23 out of 70 appeared to develop normally for the higher concentration (0.7). Do these data indicate that the rate of normal RNA synthesis is lower for cells exposed to the higher concentrations of actinomycin D?

a Find the p-value for the test.

b If you chose to use $\alpha = .05$ what is your conclusion?

10.53 How would you like to live to be 200 years old? For centuries, humankind has sought the key to the mystery of aging. What causes aging? How can aging be slowed? Studies have focused on *biomarkers*, physical or biological changes that occur at a predictable time in a person's life. The theory is that, if ways can be found to delay the occurrence of these biomarkers, human life can be extended. A key biomarker, according to scientists, is forced vital capacity (FVC), the volume of air that a person can expel after taking a deep breath. A study of 5209 men and women aged 30 to 62 showed that FVC declined, on the average, 3.8 deciliters (dl) per decade for men and 3.1 deciliters per decade for women.[10] Suppose that you wished to determine whether a physical fitness program for men and women aged 50 to 60 would delay aging; to do so, you measured the FVC for 30 men and 30 women participating in the fitness program at the beginning and end of the 50- to 60-year age interval and recorded the drop in FVC for each person. A summary of the data appears in the accompanying table.

	Men	Women
Sample size	30	30
Sample average drop in FVC (dl)	3.6	2.7
Sample standard deviation (dl)	1.1	1.2
Population mean drop in FVC	μ_1	μ_2

10. *Source:* T. Boddé, "Biomarkers of Aging: Key to a Younger Life," *Bioscience* 31(8) (1981): 566–567.