Two additional examples of the p-value may be based on Examples 8.1-3 and 8.1-4. In two-sided tests for means and proportions, the p-value is the probability of the extreme values in both directions. With the mouse data (Example 8.1-3), the p-value is

$$p\text{-value} = P(|T| \geq 0.75).$$

In Table VI in Appendix B, we see that if T has a t distribution with eight degrees of freedom, then $P(T \geq 0.706) = 0.25$. Thus, $P(|T| \geq 0.706) = 0.50$ and the p-value will be a little smaller than 0.50. In fact, $P(|T| \geq 0.75) = 0.4747$ (a probability that was found with Minitab), which is not less than $\alpha = 0.10$; hence, we do not reject H_0 at that significance level. In the example concerned with waste (Example 8.1-4), the p-value is essentially zero, since $P(T \leq -8.30) \approx 0$, where T has a t distribution with 24 degrees of freedom. Consequently, we reject H_0.

The other way of looking at tests of hypotheses is through the consideration of confidence intervals, particularly for two-sided alternatives and the corresponding tests. For example, with the mouse data (Example 8.1-3), a 90% confidence interval for the unknown mean is

$$4.3 \pm (1.86)(1.2)/\sqrt{9}, \qquad \text{or} \qquad [3.56, 5.04],$$

since $t_{0.05}(8) = 1.86$. Note that this confidence interval covers the hypothesized value $\mu = 4.0$ and we do not reject H_0: $\mu = 4.0$. If the confidence interval did not cover $\mu = 4.0$, then we would have rejected H_0: $\mu = 4.0$. Many statisticians believe that estimation is much more important than tests of hypotheses and accordingly approach statistical tests through confidence intervals. For one-sided tests, we use one-sided confidence intervals.

Exercises

8.1-1. Assume that IQ scores for a certain population are approximately $N(\mu, 100)$. To test H_0: $\mu = 110$ against the one-sided alternative hypothesis H_1: $\mu > 110$, we take a random sample of size $n = 16$ from this population and observe $\bar{x} = 113.5$.

(a) Do we accept or reject H_0 at the 5% significance level?

(b) Do we accept or reject H_0 at the 10% significance level?

(c) What is the p-value of this test?

8.1-2. Assume that the weight of cereal in a "12.6-ounce box" is $N(\mu, 0.2^2)$. The Food and Drug Association (FDA) allows only a small percentage of boxes to contain less than 12.6 ounces. We shall test the null hypothesis H_0: $\mu = 13$ against the alternative hypothesis H_1: $\mu < 13$.

(a) Use a random sample of $n = 25$ to define the test statistic and the critical region that has a significance level of $\alpha = 0.025$.

(b) If $\bar{x} = 12.9$, what is your conclusion?

(c) What is the p-value of this test?

8.1-3. Let X equal the Brinell hardness measurement of ductile iron subcritically annealed. Assume that the distribution of X is $N(\mu, 100)$. We shall test the null hypothesis H_0: $\mu = 170$ against the alternative hypothesis H_1: $\mu > 170$, using $n = 25$ observations of X.

(a) Define the test statistic and a critical region that has a significance level of $\alpha = 0.05$. Sketch a figure showing this critical region.

(b) A random sample of $n = 25$ observations of X yielded the following measurements:

170	167	174	179	179	156	163	156	187
156	183	179	174	179	170	156	187	
179	183	174	187	167	159	170	179	

Calculate the value of the test statistic and state your conclusion clearly.

(c) Give the approximate p-value of this test.

8.1-4. Let X equal the thickness of spearmint gum manufactured for vending machines. Assume that the distribution of X is $N(\mu, \sigma^2)$. The target thickness is 7.5

hundredths of an inch. We shall test the null hypothesis H_0: $\mu = 7.5$ against a two-sided alternative hypothesis, using 10 observations.

(a) Define the test statistic and critical region for an $\alpha = 0.05$ significance level. Sketch a figure illustrating this critical region.

(b) Calculate the value of the test statistic and state your decision clearly, using the following $n = 10$ thicknesses in hundredths of an inch for pieces of gum that were selected randomly from the production line:

7.65	7.60	7.65	7.70	7.55
7.55	7.40	7.40	7.50	7.50

(c) Is $\mu = 7.50$ contained in a 95% confidence interval for μ?

8.1-5. The mean birth weight of infants in the United States is $\mu = 3315$ grams. Let X be the birth weight (in grams) of a randomly selected infant in Jerusalem. Assume that the distribution of X is $N(\mu, \sigma^2)$, where μ and σ^2 are unknown. We shall test the null hypothesis H_0: $\mu = 3315$ against the alternative hypothesis H_1: $\mu < 3315$, using $n = 30$ randomly selected Jerusalem infants.

(a) Define a critical region that has a significance level of $\alpha = 0.05$.

(b) If the random sample of $n = 30$ yielded $\bar{x} = 3189$ and $s = 488$, what would be your conclusion?

(c) What is the approximate p-value of your test?

8.1-6. Let X equal the forced vital capacity (FVC) in liters for a female college student. (The FVC is the amount of air that a student can force out of her lungs.) Assume that the distribution of X is approximately $N(\mu, \sigma^2)$. Suppose it is known that $\mu = 3.4$ liters. A volleyball coach claims that the FVC of volleyball players is greater than 3.4. She plans to test her claim with a random sample of size $n = 9$.

(a) Define the null hypothesis.

(b) Define the alternative (coach's) hypothesis.

(c) Define the test statistic.

(d) Define a critical region for which $\alpha = 0.05$. Draw a figure illustrating your critical region.

(e) Calculate the value of the test statistic given that the random sample yielded the following FVCs:

3.4 3.6 3.8 3.3 3.4 3.5 3.7 3.6 3.7

(f) What is your conclusion?

(g) What is the approximate p-value of this test?

8.1-7. Vitamin B_6 is one of the vitamins in a multiple vitamin pill manufactured by a pharmaceutical company. The pills are produced with a mean of 50 mg of vitamin B_6 per pill. The company believes that there is a deterioration of 1 mg/month, so that after 3 months it expects that $\mu = 47$. A consumer group suspects that $\mu < 47$ after 3 months.

(a) Define a critical region to test H_0: $\mu = 47$ against H_1: $\mu < 47$ at an $\alpha = 0.05$ significance level based on a random sample of size $n = 20$.

(b) If the 20 pills yielded a mean of $\bar{x} = 46.94$ with a standard deviation of $s = 0.15$, what is your conclusion?

(c) What is the approximate p-value of this test?

8.1-8. A company that manufactures brackets for an automaker regularly selects brackets from the production line and performs a torque test. The goal is for mean torque to equal 125. Let X equal the torque and assume that X is $N(\mu, \sigma^2)$. We shall use a sample of size $n = 15$ to test H_0: $\mu = 125$ against a two-sided alternative hypothesis.

(a) Give the test statistic and a critical region with significance level $\alpha = 0.05$. Sketch a figure illustrating the critical region.

(b) Use the following observations to calculate the value of the test statistic and state your conclusion:

128	149	136	114	126	142	124	136
122	118	122	129	118	122	129	

8.1-9. The ornamental ground cover *Vinca minor* is spreading rapidly through the Hope College Biology Field Station because it can outcompete the small, native woody vegetation. In an attempt to discover whether *Vinca minor* utilized natural chemical weapons to inhibit the growth of the native vegetation, Hope biology students conducted an experiment in which they treated 33 sunflower seedlings with extracts taken from *Vinca minor* roots for several weeks and then measured the heights of the seedlings. Let X equal the height of one of these seedlings and assume that the distribution of X is $N(\mu, \sigma^2)$. The observed growths (in cm) were

11.5	11.8	15.7	16.1	14.1	10.5	15.2	19.0	12.8	12.4	19.2
13.5	16.5	13.5	14.4	16.7	10.9	13.0	15.1	17.1	13.3	12.4
8.5	14.3	12.9	11.1	15.0	13.3	15.8	13.5	9.3	12.2	10.3

The students also planted some control sunflower seedlings that had a mean height of 15.7 cm. We shall test the null hypothesis H_0: $\mu = 15.7$ against the alternative hypothesis H_1: $\mu < 15.7$.

could be employed to decide whether to use T or a modification of Z. However, most statisticians do not place much confidence in this test of $\sigma_X^2 = \sigma_Y^2$ and would use a modification of Z (possibly Welch's) if they suspected that the variances differed greatly. Alternatively, nonparametric methods described in Section 8.4 could be used. ■

Exercises

(In some of the exercises that follow, we must make assumptions such as the existence of normal distributions with equal variances.)

8.2-1. The botanist in Example 8.2-1 is really interested in testing for synergistic interaction. That is, given the two hormones gibberellin (GA_3) and indoleacetic acid (IAA), let X_1 and X_2 equal the growth responses (in mm) of dwarf pea stem segments to GA_3 and IAA, respectively and separately. Also, let $X = X_1 + X_2$ and let Y equal the growth response when both hormones are present. Assuming that X is $N(\mu_X, \sigma^2)$ and Y is $N(\mu_Y, \sigma^2)$, the botanist is interested in testing the hypothesis H_0: $\mu_X = \mu_Y$ against the alternative hypothesis of synergistic interaction H_1: $\mu_X < \mu_Y$.

(a) Using $n = m = 10$ observations of X and Y, define the test statistic and the critical region. Sketch a figure of the t pdf and show the critical region on your figure. Let $\alpha = 0.05$.

(b) Given $n = 10$ observations of X, namely,

2.1 2.6 2.6 3.4 2.1 1.7 2.6 2.6 2.2 1.2

and $m = 10$ observations of Y, namely,

3.5 3.9 3.0 2.3 2.1 3.1 3.6 1.8 2.9 3.3

calculate the value of the test statistic and state your conclusion. Locate the test statistic on your figure.

(c) Construct two box plots on the same figure. Does this confirm your conclusion?

8.2-2. Let X and Y denote the weights in grams of male and female common gallinules, respectively. Assume that X is $N(\mu_X, \sigma_X^2)$ and Y is $N(\mu_Y, \sigma_Y^2)$.

(a) Given $n = 16$ observations of X and $m = 13$ observations of Y, define a test statistic and a critical region for testing the null hypothesis H_0: $\mu_X = \mu_Y$ against the one-sided alternative hypothesis H_1: $\mu_X > \mu_Y$. Let $\alpha = 0.01$. (Assume that the variances are equal.)

(b) Given that $\bar{x} = 415.16$, $s_x^2 = 1356.75$, $\bar{y} = 347.40$, and $s_y^2 = 692.21$, calculate the value of the test statistic and state your conclusion.

(c) Although we assumed that $\sigma_X^2 = \sigma_Y^2$, let us say we suspect that that equality is not valid. Thus, use the test proposed by Welch.

8.2-3. Let X equal the weight in grams of a Low-Fat Strawberry Kudo and Y the weight of a Low-Fat Blueberry Kudo. Assume that the distributions of X and Y are $N(\mu_X, \sigma_X^2)$ and $N(\mu_Y, \sigma_Y^2)$, respectively. Let

21.7 21.0 21.2 20.7 20.4 21.9 20.2 21.6 20.6

be $n = 9$ observations of X, and let

21.5 20.5 20.3 21.6 21.7 21.3 23.0
21.3 18.9 20.0 20.4 20.8 20.3

be $m = 13$ observations of Y. Use these observations to answer the following questions:

(a) Test the null hypothesis H_0: $\mu_X = \mu_Y$ against a two-sided alternative hypothesis. You may select the significance level. Assume that the variances are equal.

(b) Construct and interpret box-and-whisker diagrams to support your conclusions.

8.2-4. Among the data collected for the World Health Organization air quality monitoring project is a measure of suspended particles, in $\mu g/m^3$. Let X and Y equal the concentration of suspended particles in $\mu g/m^3$ in the city centers (commercial districts), of Melbourne and Houston, respectively. Using $n = 13$ observations of X and $m = 16$ observations of Y, we shall test H_0: $\mu_X = \mu_Y$ against H_1: $\mu_X < \mu_Y$.

(a) Define the test statistic and critical region, assuming that the variances are equal. Let $\alpha = 0.05$.

(b) If $\bar{x} = 72.9$, $s_x = 25.6$, $\bar{y} = 81.7$, and $s_y = 28.3$, calculate the value of the test statistic and state your conclusion.

(c) Give bounds for the p-value of this test.

8.2-5. Some nurses in county public health conducted a survey of women who had received inadequate prenatal care. They used information from birth certificates to select mothers for the survey. The mothers selected were divided into two groups: 14 mothers who said they had five or fewer prenatal visits and 14 mothers who said they had six or more prenatal visits. Let X and Y equal the respective birth weights of the babies from these two sets of mothers, and assume that the distribution of X is $N(\mu_X, \sigma^2)$ and the distribution of Y is $N(\mu_Y, \sigma^2)$.

Sunday 5/26

(a) Define the test statistic and critical region for testing H_0: $\mu_X - \mu_Y = 0$ against H_1: $\mu_X - \mu_Y < 0$. Let $\alpha = 0.05$.

(b) Given that the observations of X were

49	108	110	82	93	114	134
114	96	52	101	114	120	116

and the observations of Y were

133	108	93	119	119	98	106
131	87	153	116	129	97	110

calculate the value of the test statistic and state your conclusion.

(c) Approximate the p-value.

(d) Construct box plots on the same figure for these two sets of data. Do the box plots support your conclusion?

8.2-6. Let X and Y equal the forces required to pull stud No. 3 and stud No. 4 out of a window that has been manufactured for an automobile. Assume that the distributions of X and Y are $N(\mu_X, \sigma_X^2)$ and $N(\mu_Y, \sigma_Y^2)$, respectively.

(a) If $m = n = 10$ observations are selected randomly, define a test statistic and a critical region for testing H_0: $\mu_X - \mu_Y = 0$ against a two-sided alternative hypothesis. Let $\alpha = 0.05$. Assume that the variances are equal.

(b) Given $n = 10$ observations of X, namely,

111 120 139 136 138 149 143 145 111 123

and $m = 10$ observations of Y, namely,

152 155 133 134 119 155 142 146 157 149

calculate the value of the test statistic and state your conclusion clearly.

(c) What is the approximate p-value of this test?

(d) Construct box plots on the same figure for these two sets of data. Do the box plots confirm your decision in part (b)?

8.2-7. Let X and Y equal the number of milligrams of tar in filtered and nonfiltered cigarettes, respectively. Assume that the distributions of X and Y are $N(\mu_X, \sigma_X^2)$ and $N(\mu_Y, \sigma_Y^2)$, respectively. We shall test the null hypothesis H_0: $\mu_X - \mu_Y = 0$ against the alternative hypothesis H_1: $\mu_X - \mu_Y < 0$, using random samples of sizes $n = 9$ and $m = 11$ observations of X and Y, respectively.

(a) Define the test statistic and a critical region that has an $\alpha = 0.01$ significance level. Sketch a figure illustrating this critical region.

(b) Given $n = 9$ observations of X, namely,

0.9 1.1 0.1 0.7 0.4 0.9 0.8 1.0 0.4

and $m = 11$ observations of Y, namely,

1.5 0.9 1.6 0.5 1.4 1.9 1.0 1.2 1.3 1.6 2.1

calculate the value of the test statistic and state your conclusion clearly. Locate the value of the test statistic on your figure.

8.2-8. Let X and Y denote the tarsus lengths of male and female grackles, respectively. Assume that X is $N(\mu_X, \sigma_X^2)$ and Y is $N(\mu_Y, \sigma_Y^2)$. Given that $n = 25$, $\bar{x} = 33.80$, $s_x^2 = 4.88$, $m = 29$, $\bar{y} = 31.66$, and $s_y^2 = 5.81$, test the null hypothesis H_0: $\mu_X = \mu_Y$ against H_1: $\mu_X > \mu_Y$ with $\alpha = 0.01$.

8.2-9. When a stream is turbid, it is not completely clear due to suspended solids in the water. The higher the turbidity, the less clear is the water. A stream was studied on 26 days, half during dry weather (say, observations of X) and the other half immediately after a significant rainfall (say, observations of Y). Assume that the distributions of X and Y are $N(\mu_X, \sigma^2)$ and $N(\mu_Y, \sigma^2)$, respectively. The following turbidities were recorded in units of NTUs (nephelometric turbidity units):

x:	2.9	14.9	1.0	12.6	9.4	7.6	3.6
	3.1	2.7	4.8	3.4	7.1	7.2	
y:	7.8	4.2	2.4	12.9	17.3	10.4	5.9
	4.9	5.1	8.4	10.8	23.4	9.7	

(a) Test the null hypothesis H_0: $\mu_X = \mu_Y$ against H_1: $\mu_X < \mu_Y$. Give bounds for the p-value and state your conclusion.

(b) Draw box-and-whisker diagrams on the same graph. Does this figure confirm your answer?

8.2-10. Plants convert carbon dioxide (CO_2) in the atmosphere, along with water and energy from sunlight, into the energy they need for growth and reproduction. Experiments were performed under normal atmospheric air conditions and in air with enriched CO_2 concentrations to determine the effect on plant growth. The plants were given the same amount of water and light for a four-week period. The following table gives the plant growths in grams:

Normal Air	4.67	4.21	2.18	3.91	4.09	5.24	2.94	4.71
	4.04	5.79	3.80	4.38				
Enriched Air	5.04	4.52	6.18	7.01	4.36	1.81	6.22	5.70

On the basis of these data, determine whether CO_2-enriched atmosphere increases plant growth.

8.2-11. Let X equal the fill weight in April and Y the fill weight in June for an 8-pound box of bleach. We shall test