10 Two-Sample Tests

Learning Objectives

In this chapter, you learn how to use hypothesis testing for comparing the difference between:

- The means of two independent populations
- The means of two related populations
- The proportions of two independent populations
- The variances of two independent populations

USING STATISTICS

@ BLK Beverages

Does the type of display used in a supermarket affect the sales of products? As the regional sales manager for BLK Beverages, you want to compare the sales volume of BLK Cola when the product is placed in the normal shelf location to the sales volume when the product is featured in a special end-aisle display. To test the effectiveness of the end-aisle displays, you select 20 stores from the Food Pride supermarket chain that all experience similar storewide sales volumes. You then randomly assign 10 of the 20 stores to sample 1 and 10 stores to sample 2. The managers of the 10 stores in sample 1 place the BLK Cola in the normal shelf location, alongside the other cola products. The 10 stores in sample 2 use the special end-aisle promotional display. At the end of one week, the sales of BLK Cola are recorded. How can you determine whether sales of BLK Cola using the end-aisle displays are the same as those when the cola is placed in the normal shelf location? How can you decide if the variability in BLK Cola sales from store to store is the same for the two types of displays? How could you use the answers to these questions to improve sales of BLK Cola?

Hypothesis testing provides a *confirmatory* approach to data analysis. In Chapter 9, you learned a variety of commonly used hypothesis-testing procedures that relate to a single sample of data selected from a single population. In this chapter, you learn how to extend hypothesis testing to **two-sample tests** that compare statistics from samples of data selected from two populations. One such test for the BLK Beverages scenario would be "Are the mean weekly sales of BLK Cola when using the normal shelf location (one population) equal to the mean weekly sales of BLK Cola when using an end-aisle display (a second population)?"

10.1 Comparing the Means of Two Independent Populations

Worksheet data for samples taken from two independent populations can be stored either in stacked or unstacked format, as discussed in Section 2.3. Examples throughout this chapter use unstacked data, although by using the techniques discussed in either Section EG2.3 (for Excel) or MG2.3 (for Minitab), you can rearrange unstacked data as stacked data or rearrange stacked data as unstacked data.

In Sections 8.1 and 9.1, you learned that in almost all cases, you would not know the population standard deviation of the population under study. Likewise, when you take a random sample from each of two independent populations, you almost always do not know the standard deviations of either population. However, you also need to know whether you can assume that the variances in the two populations are equal because the method you use to compare the means of each population depends on whether you can assume that the variances of the two populations are equal.

Pooled-Variance *t* Test for the Difference Between Two Means

If you assume that the random samples are independently selected from two populations and that the populations are normally distributed and have equal variances, you can use a **pooled-variance *t* test** to determine whether there is a significant difference between the means of the two populations. If the populations are not normally distributed, the pooled-variance t test can still be used if the sample sizes are large enough (typically ≥ 30 for each sample[1]).

Using subscripts to distinguish between the population mean of the first population, μ_1, and the population mean of the second population, μ_2, the null hypothesis of no difference in the means of two independent populations can be stated as

$$H_0: \mu_1 = \mu_2 \text{ or } \mu_1 - \mu_2 = 0$$

and the alternative hypothesis, that the means are not the same, can be stated as

$$H_1: \mu_1 \neq \mu_2 \text{ or } \mu_1 - \mu_2 \neq 0$$

To test the null hypothesis, use the pooled-variance t test statistic t_{STAT} shown in Equation (10.1). The pooled-variance t test gets its name from the fact that the test statistic pools, or combines, the two sample variances S_1^2 and S_2^2 to compute S_p^2, the best estimate of the variance common to both populations, under the assumption that the two population variances are equal.[2]

[1]Review the Section 7.4 discussion about the Central Limit Theorem on page 264 to understand more about "large enough" sample sizes.

[2]When the two sample sizes are equal (i.e., $n_1 = n_2$), the equation for the pooled variance can be simplified to

$$S_p^2 = \frac{S_1^2 + S_2^2}{2}$$

POOLED-VARIANCE *t* TEST FOR THE DIFFERENCE BETWEEN TWO MEANS

$$t_{STAT} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \textbf{(10.1)}$$

where

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$

and

S_p^2 = pooled variance

$\bar{X}_1$ = mean of the sample taken from population 1

S_1^2 = variance of the sample taken from population 1

n_1 = size of the sample taken from population 1

$\bar{X}_2$ = mean of the sample taken from population 2

S_2^2 = variance of the sample taken from population 2

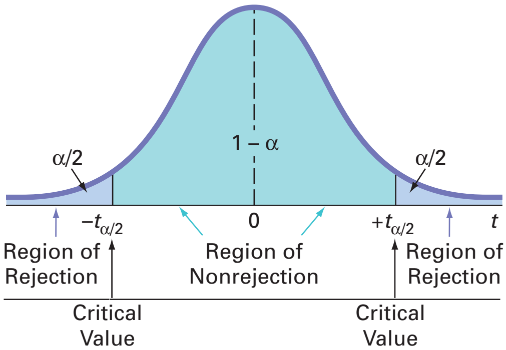
n_2 = size of the sample taken from population 2

The t_{STAT} test statistic follows a t distribution with $n_1 + n_2 - 2$ degrees of freedom.

For a given level of significance, α, in a two-tail test, you reject the null hypothesis if the computed t_{STAT} test statistic is greater than the upper-tail critical value from the t distribution or if the computed t_{STAT} test statistic is less than the lower-tail critical value from the t distribution. Figure 10.1 displays the regions of rejection.

FIGURE 10.1

Regions of rejection and nonrejection for the pooled-variance t test for the difference between the means (two-tail test)



In a one-tail test in which the rejection region is in the lower tail, you reject the null hypothesis if the computed t_{STAT} test statistic is less than the lower-tail critical value from the t distribution. In a one-tail test in which the rejection region is in the upper tail, you reject the null hypothesis if the computed t_{STAT} test statistic is greater than the upper-tail critical value from the t distribution.

To demonstrate the pooled-variance t test, return to the BLK Beverages scenario on page 365. You define the business objective as determining whether the mean weekly sales of BLK Cola are the same when using a normal shelf location and when using an end-aisle display. There are two populations of interest. The first population is the set of all possible weekly sales of BLK Cola if all the Food Pride Supermarkets used the normal shelf location. The second population is the set of all possible weekly sales of BLK Cola if all the Food Pride Supermarkets used the end-aisle displays. You collect the data from a sample of 10 Food Pride Supermarkets that have been assigned a normal shelf location and another sample of 10 Food Pride Supermarkets that have been assigned an end-aisle display. You organize and store the results in Cola. Table 10.1 contains the BLK Cola sales (in number of cases) for the two samples.

TABLE 10.1

Comparing BLK Cola Weekly Sales from Two Different Display Locations (in number of cases)

Display Location									
Normal					**End-Aisle**				
22	34	52	62	30	52	71	76	54	67
40	64	84	56	59	83	66	90	77	84

The null and alternative hypotheses are

$$H_0: \mu_1 = \mu_2 \text{ or } \mu_1 - \mu_2 = 0$$

$$H_1: \mu_1 \neq \mu_2 \text{ or } \mu_1 - \mu_2 \neq 0$$

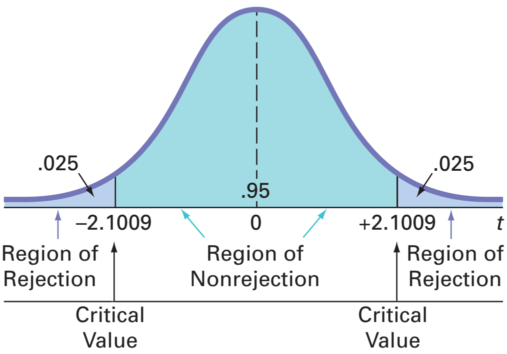
Assuming that the samples are from normal populations having equal variances, you can use the pooled-variance t test. The t_{STAT} test statistic follows a t distribution with $10 + 10 - 2 = 18$

degrees of freedom. Using an $\alpha = 0.05$ level of significance, you divide the rejection region into the two tails for this two-tail test (i.e., two equal parts of 0.025 each). Table E.3 shows that the critical values for this two-tail test are $+2.1009$ and -2.1009. As shown in Figure 10.2, the decision rule is

$$\text{Reject } H_0 \text{ if } t_{STAT} > +2.1009$$
$$\text{or if } t_{STAT} < -2.1009;$$
$$\text{otherwise do not reject } H_0.$$

FIGURE 10.2

Two-tail test of hypothesis for the difference between the means at the 0.05 level of significance with 18 degrees of freedom



From Figure 10.3, the computed t_{STAT} test statistic for this test is -3.0446 and the p-value is 0.0070.

FIGURE 10.3

Excel and Minitab results for the pooled-variance t test for the BLK Cola display locations

	A	B	
1	Pooled-Variance *t* Test for the Difference Between Two Means		
2	(assumes equal population variances)		
3	Data		
4	Hypothesized Difference	0	
5	Level of Significance	0.05	
6	Population 1 Sample		
7	Sample Size	10	=COUNT(DATACOPY!$A:$A)
8	Sample Mean	50.3	=AVERAGE(DATACOPY!$A:$A)
9	Sample Standard Deviation	18.7264	=STDEV(DATACOPY!$A:$A)
10	Population 2 Sample		
11	Sample Size	10	=COUNT(DATACOPY!$B:$B)
12	Sample Mean	72	=AVERAGE(DATACOPY!$B:$B)
13	Sample Standard Deviation	12.5433	=STDEV(DATACOPY!$B:$B)
14			
15	Intermediate Calculations		
16	Population 1 Sample Degrees of Freedom	9	=B7 - 1
17	Population 2 Sample Degrees of Freedom	9	=B11 - 1
18	Total Degrees of Freedom	18	=B16 + B17
19	Pooled Variance	254.0056	=((B16 * B9^2) + (B17 * B13^2))/B18
20	Standard Error	7.1275	=SQRT(B19 * (1/B7 + 1/B11))
21	Difference in Sample Means	-21.7	=B8 - B12
22	*t* Test Statistic	-3.0446	=(B21 - B4)/B20
23			
24	Two-Tail Test		
25	Lower Critical Value	-2.1009	=-TINV(B5, B18)
26	Upper Critical Value	2.1009	=TINV(B5, B18)
27	*p*-Value	0.0070	=TDIST(ABS(B22), B18, 2)
28	Reject the null hypothesis		=IF(B27 < B5, "Reject the null hypothesis", "Do not reject the null hypothesis")

Two-Sample T-Test and CI: Normal, End-Aisle

```
Two-sample T for Normal vs End-Aisle

            N  Mean  StDev  SE Mean
Normal     10  50.3   18.7      5.9
End-Aisle  10  72.0   12.5      4.0

Difference = mu (Normal) - mu (End-Aisle)
Estimate for difference:  -21.70
95% CI for difference:  (-36.67, -6.73)
T-Test of difference = 0 (vs not =): T-Value = -3.04  P-Value = 0.007  DF = 18
Both use Pooled StDev = 15.9376
```

Using Equation (10.1) on page 366 and the descriptive statistics provided in Figure 10.3,

$$t_{STAT} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

where

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$
$$= \frac{9(18.7264)^2 + 9(12.5433)^2}{9 + 9} = 254.0056$$

Therefore,

$$t_{STAT} = \frac{(50.3 - 72.0) - 0.0}{\sqrt{254.0056\left(\frac{1}{10} + \frac{1}{10}\right)}} = \frac{-21.7}{\sqrt{50.801}} = -3.0446$$

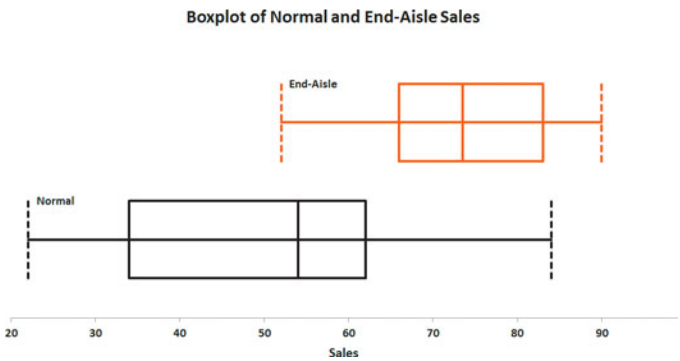
You reject the null hypothesis because $t_{STAT} = -3.0446 < -2.1009$ and the p-value is 0.0070. In other words, the probability that $t_{STAT} > 3.0446$ or $t_{STAT} < -3.0446$ is equal to 0.0070. This p-value indicates that if the population means are equal, the probability of observing a difference this large or larger in the two sample means is only 0.0070. Because the p-value is less than $\alpha = 0.05$, there is sufficient evidence to reject the null hypothesis. You can conclude that the mean sales are different for the normal shelf location and the end-aisle location. Based on these results, the sales are lower for the normal location (i.e., higher for the end-aisle location).

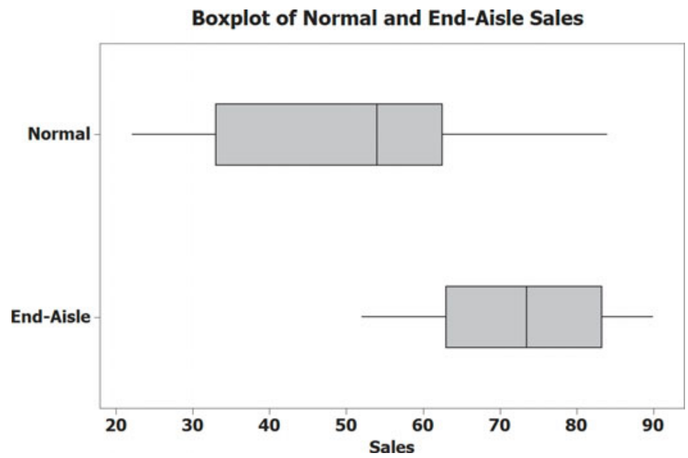
In testing for the difference between the means, you assume that the populations are normally distributed, with equal variances. For situations in which the two populations have equal variances, the pooled-variance t test is **robust** (i.e., not sensitive) to moderate departures from the assumption of normality, provided that the sample sizes are large. In such situations, you can use the pooled-variance t test without serious effects on its power. However, if you cannot assume that both populations are normally distributed, you have two choices. You can use a nonparametric procedure, such as the Wilcoxon rank sum test (see Section 12.6), that does not depend on the assumption of normality for the two populations, or you can use a normalizing transformation (see reference 6) on each of the outcomes and then use the pooled-variance t test.

To check the assumption of normality in each of the two populations, construct the boxplot of the sales for the two display locations shown in Figure 10.4. For these two small samples, there appears to be only moderate departure from normality, so the assumption of normality needed for the t test is not seriously violated.

FIGURE 10.4

Excel and Minitab boxplots of the sales for the two display locations





Example 10.1 provides another application of the pooled-variance t test.

EXAMPLE 10.1

Testing for the Difference in the Mean Delivery Times

You and some friends have decided to test the validity of an advertisement by a local pizza restaurant, which says it delivers to the dormitories faster than a local branch of a national chain. Both the local pizza restaurant and national chain are located across the street from your college campus. You define the variable of interest as the delivery time, in minutes, from the time the pizza is ordered to when it is delivered. You collect the data by ordering 10 pizzas from the local pizza restaurant and 10 pizzas from the national chain at different times. You organize and store the data in **PizzaTime**. Table 10.2 shows the delivery times.

TABLE 10.2

Delivery Times (in minutes) for Local Pizza Restaurant and National Pizza Chain

Local		Chain	
16.8	18.1	22.0	19.5
11.7	14.1	15.2	17.0
15.6	21.8	18.7	19.5
16.7	13.9	15.6	16.5
17.5	20.8	20.8	24.0

At the 0.05 level of significance, is there evidence that the mean delivery time for the local pizza restaurant is less than the mean delivery time for the national pizza chain?

SOLUTION Because you want to know whether the mean is *lower* for the local pizza restaurant than for the national pizza chain, you have a one-tail test with the following null and alternative hypotheses:

H_0: $\mu_1 \geq \mu_2$ (The mean delivery time for the local pizza restaurant is equal to or greater than the mean delivery time for the national pizza chain.)

H_1: $\mu_1 < \mu_2$ (The mean delivery time for the local pizza restaurant is less than the mean delivery time for the national pizza chain.)

Figure 10.5 displays the results for the pooled-variance *t* test for these data.

FIGURE 10.5

Excel and Minitab results for the pooled-variance *t* test for the pizza delivery time data

	A	B
1	Pooled-Variance *t* Test for the Difference Between Two Means	
2	(assumes equal population variances)	
3	Data	
4	Hypothesized Difference	0
5	Level of Significance	0.05
6	Population 1 Sample	
7	Sample Size	10
8	Sample Mean	16.7
9	Sample Standard Deviation	3.0955
10	Population 2 Sample	
11	Sample Size	10
12	Sample Mean	18.88
13	Sample Standard Deviation	2.8662
14		
15	Intermediate Calculations	
16	Population 1 Sample Degrees of Freedom	9
17	Population 2 Sample Degrees of Freedom	9
18	Total Degrees of Freedom	18
19	Pooled Variance	8.8987
20	Standard Error	1.3341
21	Difference in Sample Means	-2.18
22	*t* Test Statistic	-1.6341
23		
24	Lower-Tail Test	
25	Lower Critical Value	-1.7341
26	*p*-Value	0.0598
27	Do not reject the null hypothesis	

Two-Sample T-Test and CI: Local, Chain

```
Two-sample T for Local vs Chain

        N   Mean  StDev  SE Mean
Local  10  16.70   3.10     0.98
Chain  10  18.88   2.87     0.91

Difference = mu (Local) - mu (Chain)
Estimate for difference:  -2.18
95% upper bound for difference:  0.13
T-Test of difference = 0 (vs <): T-Value = -1.63  P-Value = 0.060  DF = 18
Both use Pooled StDev = 2.9831
```

To illustrate the computations, using Equation (10.1) on page 366,

$$t_{STAT} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}$$

$$= \frac{9(3.0955)^2 + 9(2.8662)^2}{9 + 9} = 8.8987$$

Therefore,

$$t_{STAT} = \frac{(16.7 - 18.88) - 0.0}{\sqrt{8.8987\left(\frac{1}{10} + \frac{1}{10}\right)}} = \frac{-2.18}{\sqrt{1.7797}} = -1.6341$$

You do not reject the null hypothesis because $t_{STAT} = -1.6341 > -1.7341$. The p-value (as computed in Figure 10.5) is 0.0598. This p-value indicates that the probability that $t_{STAT} < -1.6341$ is equal to 0.0598. In other words, if the population means are equal, the probability that the sample mean delivery time for the local pizza restaurant is at least 2.18 minutes faster than the national chain is 0.0598. Because the p-value is greater than $\alpha = 0.05$, there is insufficient evidence to reject the null hypothesis. Based on these results, there is insufficient evidence for the local pizza restaurant to make the advertising claim that it has a faster delivery time.

Confidence Interval Estimate for the Difference Between Two Means

Instead of, or in addition to, testing for the difference in the means of two independent populations, you can use Equation (10.2) to develop a confidence interval estimate of the difference in the means.

CONFIDENCE INTERVAL ESTIMATE FOR THE DIFFERENCE IN THE MEANS OF TWO INDEPENDENT POPULATIONS

$$(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

or

$$(\bar{X}_1 - \bar{X}_2) - t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \le \mu_1 - \mu_2 \le (\bar{X}_1 - \bar{X}_2) + t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \quad \textbf{(10.2)}$$

where $t_{\alpha/2}$ is the critical value of the t distribution, with $n_1 + n_2 - 2$ degrees of freedom, for an area of $\alpha/2$ in the upper tail.

For the sample statistics pertaining to the two aisle locations reported in Figure 10.3 on page 368, using 95% confidence, and Equation (10.2),

$\bar{X}_1 = 50.3, n_1 = 10, \bar{X}_2 = 72.0, n_2 = 10, S_p^2 = 254.0056$, and with $10 + 10 - 2$

$= 18$ degrees of freedom, $t_{0.025} = 2.1009$

$$(50.3 - 72.0) \pm (2.1009)\sqrt{254.0056\left(\frac{1}{10} + \frac{1}{10}\right)}$$

$$-21.7 \pm (2.1009)(7.1275)$$

$$-21.7 \pm 14.97$$

$$-36.67 \le \mu_1 - \mu_2 \le -6.73$$

Therefore, you are 95% confident that the difference in mean sales between the normal aisle location and the end-aisle location is between −36.67 cases of cola and −6.73 cases of cola. In other words, the end-aisle location sells, on average, 6.73 to 36.67 cases more than the normal aisle location. From a hypothesis-testing perspective, because the interval does not include zero, you reject the null hypothesis of no difference between the means of the two populations.

t Test for the Difference Between Two Means, Assuming Unequal Variances

If you cannot make the assumption that the two independent populations have equal variances, you cannot pool the two sample variances into the common estimate S_p^2 and therefore cannot use the pooled-variance t test. Instead, you use the **separate-variance *t* test** developed by Satterthwaite (see reference 5). Equation (10.3) defines the test statistic for the separate-variance t test.

SEPARATE-VARIANCE t TEST FOR THE DIFFERENCE BETWEEN TWO MEANS

$$t_{STAT} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}} \quad \textbf{(10.3)}$$

where

$\bar{X}_1$ = mean of the sample taken from population 1

S_1^2 = variance of the sample taken from population 1

n_1 = size of the sample taken from population 1

$\bar{X}_2$ = mean of the sample taken from population 2

S_2^2 = variance of the sample taken from population 2

n_2 = size of the sample taken from population 2

The separate-variance t test statistic approximately follows a t distribution with V degrees of freedom equal to the integer portion of the following computation.

COMPUTING DEGREES OF FREEDOM IN THE SEPARATE-VARIANCE t TEST

$$V = \frac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{S_2^2}{n_2}\right)^2}{n_2 - 1}} \quad \textbf{(10.4)}$$

For a given level of significance α, you reject the null hypothesis if the computed t test statistic is greater than the upper-tail critical value $t_{a/2}$ from the t distribution with V degrees of freedom or if the computed t test statistic is less than the lower-tail critical value $-t_{a/2}$ from the t distribution with V degrees of freedom. Thus, the decision rule is

Reject H_0 if $t > t_{a/2}$

or if $t < -t_{a/2}$;

otherwise, do not reject H_0.

Recall the Using Statistics scenario for this chapter that concerned the display location of BLK Cola. Using Equation (10.4), the separate-variance t test statistic t_{STAT} is approximated by a t distribution with $V = 15$ degrees of freedom, the integer portion of the following computation:

$$V = \frac{\left(\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}\right)^2}{\dfrac{\left(\dfrac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \dfrac{\left(\dfrac{S_2^2}{n_2}\right)^2}{n_2 - 1}}$$

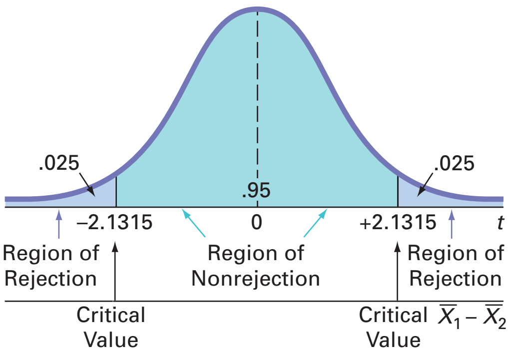
$$= \frac{\left(\frac{350.6778}{10} + \frac{157.3333}{10}\right)^2}{\frac{\left(\frac{350.6778}{10}\right)^2}{9} + \frac{\left(\frac{157.3333}{10}\right)^2}{9}} = 15.72$$

Using $\alpha = 0.05$, the upper and lower critical values for this two-tail test found in Table E.3 are +2.1315 and −2.1315. As depicted in Figure 10.6, the decision rule is

$$\text{Reject } H_0 \text{ if } t_{STAT} > +2.1315$$
$$\text{or if } t_{STAT} < -2.1315;$$
$$\text{otherwise do not reject } H_0.$$

FIGURE 10.6

Two-tail test of hypothesis for the difference between the means at the 0.05 level of significance with 15 degrees of freedom



Using Equation (10.3) on page 372 and the descriptive statistics provided in Figure 10.3,

$$t_{STAT} = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$= \frac{50.3 - 72}{\sqrt{\left(\frac{350.6778}{10} + \frac{157.3333}{10}\right)}} = \frac{-21.7}{\sqrt{50.801}} = -3.04$$

Using a 0.05 level of significance, you reject the null hypothesis because $t = -3.04 < -2.1315$. Figure 10.7 on page 374 displays the separate-variance t test results for the display location data.

In Figure 10.7, the test statistic $t_{STAT} = -3.0446$ and the p-value is $0.0082 < 0.05$. Thus, the results for the separate-variance t test are almost exactly the same as those of the pooled-variance t test. The assumption of equality of population variances had no appreciable effect on the results. Sometimes, however, the results from the pooled-variance and separate-variance t tests conflict because the assumption of equal variances is violated. Therefore, it is important that you evaluate the assumptions and use those results as a guide in selecting a test procedure. In Section 10.4, the F test for the ratio of two variances is used to determine whether there is evidence of a difference in the two population variances. The results of that test can help you determine which of the t tests—pooled-variance or separate-variance—is more appropriate.

FIGURE 10.7

Excel and Minitab results for the separate-variance *t* test for the display location data

	A	B	
1	Separate-Variances *t* Test for the Difference Between Two Means		
2	(assumes unequal population variances)		
3	Data		
4	Hypothesized Difference	0	
5	Level of Significance	0.05	
6	Population 1 Sample		
7	Sample Size	10	=COUNT(DATACOPY!$A:$A)
8	Sample Mean	50.3	=AVERAGE(DATACOPY!$A:$A)
9	Sample Standard Deviation	18.7264	=STDEV(DATACOPY!$A:$A)
10	Population 2 Sample		
11	Sample Size	10	=COUNT(DATACOPY!$B:$B)
12	Sample Mean	72	=AVERAGE(DATACOPY!$B:$B)
13	Sample Standard Deviation	12.5433	=STDEV(DATACOPY!$B:$B)
14			
15	Intermediate Calculations		
16	Numerator of Degrees of Freedom	2580.7529	=(E18 + E19)^2
17	Denominator of Degrees of Freedom	164.1430	=(E18^2)/(B7 - 1) + (E19^2)/(B11 - 1)
18	Total Degrees of Freedom	15.7226	=B16/B17
19	Degrees of Freedom	15	=INT(B18)
20	Standard Error	7.1275	=SQRT(E18 + E19)
21	Difference in Sample Means	-21.7	=B8 - B12
22	Separate-Variance *t* Test Statistic	-3.0446	=B21/B20
23			
24	Two-Tail Test		
25	Lower Critical Value	-2.1314	=-TINV(B5, B19)
26	Upper Critical Value	2.1314	=TINV(B5, B19)
27	*p*-Value	0.0082	=TDIST(ABS(B22), B19, 2)
28	Reject the null hypothesis		=IF(B27 < B5, "Reject the null hypothesis", "Do not reject the null hypothesis")

Not shown
The Calculations Area in cell range D15:E22

Two-Sample T-Test and CI: Normal, End-Aisle

```
Two-sample T for Normal vs End-Aisle

            N  Mean  StDev  SE Mean
Normal     10  50.3   18.7      5.9
End-Aisle  10  72.0   12.5      4.0

Difference = mu (Normal) - mu (End-Aisle)
Estimate for difference:  -21.70
95% CI for difference:  (-36.89, -6.51)
T-Test of difference = 0 (vs not =): T-Value = -3.04  P-Value = 0.008  DF = 15
```

THINK ABOUT THIS “This Call May Be Monitored ... ”

When talking with a customer service representative by phone, you may have heard a “This call may be monitored ... ” message. Typically, the message explains that the monitoring is for “quality assurance purposes,” but do companies really monitor your calls to improve quality?

From a past student, we’ve discovered that at least one large financial services company really does monitor the quality of calls. This student was asked to develop an improved training program for a call center that was hiring people to answer phone calls customers make about outstanding loans. For feedback and evaluation, she planned to randomly select phone calls received by each new employee and rate the employee on 10 aspects of the call, including whether the employee maintained a pleasant tone with the customer.

Who You Gonna Call?

The former student presented her plan to her boss for approval, but her boss, quoting a famous statistician, said, “In God we trust, all others must bring data.” *Her boss wanted proof that her new training program would improve customer service.* Faced with this request, who would you call? She called her business statistics professor, one of the co-authors of this book. “Hey, Professor, you’ll never believe why I called. I work for a large company, and in the project I am currently working on, I have to put some of the statistics you taught us to work! Can you help?” Together they formulated this test:

- Randomly assign the 60 most recent hires to two training programs. Assign half to the preexisting training program and the other half to the new training program.
- At the end of the first month, compare the mean score for the 30 employees in the new training program against the mean score for the 30 employees in the preexisting training program.

She listened as her professor explained, “What you are trying to show is that the mean score from the new training program is higher than the mean score from the current program. You can make the null hypothesis that the means are equal and see if you can reject it in favor of the alternative that the mean score from the new program is higher.”

“Or, as you used to say, ‘if the *p*-value is low, H_0 must go!’—yes, I do remember!” she replied. Her professor chuckled and added, “If you can reject H_0, you will have the evidence to present to your boss.” She thanked him for his help and got back to work, with the newfound confidence that she would be able to successfully apply the *t* test that compares the means of two independent populations.

Problems for Section 10.1

LEARNING THE BASICS

10.1 If you have samples of $n_1 = 12$ and $n_2 = 15$, in performing the pooled-variance *t* test, how many degrees of freedom do you have?

10.2 Assume that you have a sample of $n_1 = 8$, with the sample mean $\bar{X}_1 = 42$, and a sample standard deviation $S_1 = 4$, and you have an independent sample of $n_2 = 15$ from another population with a sample mean of $\bar{X}_2 = 34$ and a sample standard deviation $S_2 = 5$.

a. What is the value of the pooled-variance t_{STAT} test statistic for testing $H_0: \mu_1 = \mu_2$?

b. In finding the critical value, how many degrees of freedom are there?

c. Using the level of significance $\alpha = 0.01$, what is the critical value for a one-tail test of the hypothesis $H_0: \mu_1 \leq \mu_2$ against the alternative, $H_1: \mu_1 > \mu_2$?
d. What is your statistical decision?

10.3 What assumptions about the two populations are necessary in Problem 10.2?

10.4 Referring to Problem 10.2, construct a 95% confidence interval estimate of the population mean difference between μ_1 and μ_2.

10.5 Referring to Problem 10.2, if $n_1 = 5$ and $n_2 = 4$, how many degrees of freedom do you have?

10.6 Referring to Problem 10.2, if $n_1 = 5$ and $n_2 = 4$, at the 0.01 level of significance, is there evidence that $\mu_1 > \mu_2$?

APPLYING THE CONCEPTS

10.7 According to a recent study, when shopping online for luxury goods, men spend a mean of $2,401, whereas women spend a mean of $1,527. (Data extracted from R. A. Smith, "Fashion Online: Retailers Tackle the Gender Gap," *The Wall Street Journal*, March 13, 2008, pp. D1, D10.) Suppose that the study was based on a sample of 600 men and 700 females, and the standard deviation of the amount spent was $1,200 for men and $1,000 for women.
a. State the null and alternative hypothesis if you want to determine whether the mean amount spent is higher for men than for women.
b. In the context of this study, what is the meaning of the Type I error?
c. In the context of this study, what is the meaning of the Type II error?
d. At the 0.01 level of significance, is there evidence that the mean amount spent is higher for men than for women?

10.8 A recent study ("Snack Ads Spur Children to Eat More," *The New York Times*, July 20, 2009, p. B3) found that children who watched a cartoon with food advertising ate, on average, 28.5 grams of Goldfish crackers as compared to an average of 19.7 grams of Goldfish crackers for children who watched a cartoon without food advertising. Although there were 118 children in the study, neither the sample size in each group nor the sample standard deviations were reported. Suppose that there were 59 children in each group, and the sample standard deviation for those children who watched the food ad was 8.6 grams and the sample standard deviation for those children who did not watch the food ad was 7.9 grams.
a. Assuming that the population variances are equal and $\alpha = 0.05$, is there evidence that the mean amount of Goldfish crackers eaten was significantly higher for the children who watched food ads?
b. Assuming that the population variances are equal, construct a 95% confidence interval estimate of the difference between the mean amount of Goldfish crackers eaten by the children who watched and did not watch the food ad.
c. Compare the results of (a) and (b) and discuss.

10.9 A problem with a telephone line that prevents a customer from receiving or making calls is upsetting to both the customer and the telephone company. The file **Phone** contains samples of 20 problems reported to two different offices of a telephone company and the time to clear these problems (in minutes) from the customers' lines:

Central Office I Time to Clear Problems (minutes)

1.48	1.75	0.78	2.85	0.52	1.60	4.15	3.97	1.48	3.10
1.02	0.53	0.93	1.60	0.80	1.05	6.32	3.93	5.45	0.97

Central Office II Time to Clear Problems (minutes)

7.55	3.75	0.10	1.10	0.60	0.52	3.30	2.10	0.58	4.02
3.75	0.65	1.92	0.60	1.53	4.23	0.08	1.48	1.65	0.72

a. Assuming that the population variances from both offices are equal, is there evidence of a difference in the mean waiting time between the two offices? (Use $\alpha = 0.05$.)
b. Find the p-value in (a) and interpret its meaning.
c. What other assumption is necessary in (a)?
d. Assuming that the population variances from both offices are equal, construct and interpret a 95% confidence interval estimate of the difference between the population means in the two offices.

SELF Test **10.10** The Computer Anxiety Rating Scale (CARS) measures an individual's level of computer anxiety, on a scale from 20 (no anxiety) to 100 (highest level of anxiety). Researchers at Miami University administered CARS to 172 business students. One of the objectives of the study was to determine whether there is a difference in the level of computer anxiety experienced by female and male business students. They found the following:

	Males	Females
$\bar{X}$	40.26	36.85
S	13.35	9.42
n	100	72

Source: Data extracted from T. Broome and D. Havelka, "Determinants of Computer Anxiety in Business Students," *The Review of Business Information Systems,* Spring 2002, 6(2), pp. 9–16.

a. At the 0.05 level of significance, is there evidence of a difference in the mean computer anxiety experienced by female and male business students?
b. Determine the p-value and interpret its meaning.
c. What assumptions do you have to make about the two populations in order to justify the use of the t test?

10.11 An important feature of digital cameras is battery life, the number of shots that can be taken before the battery needs to be recharged. The file **DigitalCameras** contains the battery life of 29 subcompact cameras and 16 compact cameras. (Data extracted from "Digital Cameras," *Consumer Reports*, July 2009, pp. 28–29.)

a. Assuming that the population variances from both types of digital cameras are equal, is there evidence of a difference in the mean battery life between the two types of digital cameras ($\alpha = 0.05$)?
b. Determine the p-value in (a) and interpret its meaning.
c. Assuming that the population variances from both types of digital cameras are equal, construct and interpret a 95% confidence interval estimate of the difference between the population mean battery life of the two types of digital cameras.

10.12 A bank with a branch located in a commercial district of a city has the business objective of developing an improved process for serving customers during the noon-to-1 P.M. lunch period. Management decides to first study the waiting time in the current process. The waiting time is defined as the time that elapses from when the customer enters the line until he or she reaches the teller window. Data are collected from a random sample of 15 customers, and the results (in minutes) are as follows (and stored in **Bank1**):

4.21 5.55 3.02 5.13 4.77 2.34 3.54 3.20
4.50 6.10 0.38 5.12 6.46 6.19 3.79

Suppose that another branch, located in a residential area, is also concerned with improving the process of serving customers in the noon-to-1 P.M. lunch period. Data are collected from a random sample of 15 customers, and the results are as follows (and stored in **Bank2**):

9.66 5.90 8.02 5.79 8.73 3.82 8.01 8.35
10.49 6.68 5.64 4.08 6.17 9.91 5.47

a. Assuming that the population variances from both banks are equal, is there evidence of a difference in the mean waiting time between the two branches? (Use $\alpha = 0.05$.)
b. Determine the p-value in (a) and interpret its meaning.
c. In addition to equal variances, what other assumption is necessary in (a)?
d. Construct and interpret a 95% confidence interval estimate of the difference between the population means in the two branches.

10.13 Repeat Problem 10.12 (a), assuming that the population variances in the two branches are not equal. Compare the results with those of Problem 10.12 (a).

10.14 In intaglio printing, a design or figure is carved beneath the surface of hard metal or stone. The business objective of an intaglio printing company is to determine whether there are differences in the mean surface hardness of steel plates, based on two different surface conditions—untreated and treated by lightly polishing with emery paper. An experiment is designed in which 40 steel plates are randomly assigned—20 plates are untreated and 20 plates are treated. The results of the experiment (stored in **Intaglio**) are as follows:

Untreated		Treated	
164.368	177.135	158.239	150.226
159.018	163.903	138.216	155.620
153.871	167.802	168.006	151.233
165.096	160.818	149.654	158.653
157.184	167.433	145.456	151.204
154.496	163.538	168.178	150.869
160.920	164.525	154.321	161.657
164.917	171.230	162.763	157.016
169.091	174.964	161.020	156.670
175.276	166.311	167.706	147.920

a. Assuming that the population variances from both conditions are equal, is there evidence of a difference in the mean surface hardness between untreated and treated steel plates? (Use $\alpha = 0.05$.)
b. Determine the p-value in (a) and interpret its meaning.
c. In addition to equal variances, what other assumption is necessary in (a)?
d. Construct and interpret a 95% confidence interval estimate of the difference between the population means from treated and untreated steel plates.

10.15 Repeat Problem 10.14 (a), assuming that the population variances from untreated and treated steel plates are not equal. Compare the results with those of Problem 10.14 (a).

10.16 Do young children use cell phones? Apparently so, according to a recent study (A. Ross, "Message to Santa; Kids Want a Phone," *Palm Beach Post*, December 16, 2008, pp. 1A, 4A), which stated that cell phone users under 12 years of age averaged 137 calls per month as compared to 231 calls per month for cell phone users 13 to 17 years of age. No sample sizes were reported. Suppose that the results were based on samples of 50 cell phone users in each group and that the sample standard deviation for cell phone users under 12 years of age was 51.7 calls per month and the sample standard deviation for cell phone users 13 to 17 years of age was 67.6 calls per month.

a. Assuming that the variances in the populations of cell phone users are equal, is there evidence of a difference in the mean cell phone usage between cell phone users under 12 years of age and cell phone users 13 to 17 years of age? (Use a 0.05 level of significance.)
b. In addition to equal variances, what other assumption is necessary in (a)?

10.17 Nondestructive evaluation is a method that is used to describe the properties of components or materials without causing any permanent physical change to the units. It includes the determination of properties of materials and the classification of flaws by size, shape, type, and location. This method is most effective for detecting surface flaws and characterizing surface properties of electrically conductive materials. Data were collected that classified each component as having a flaw or not, based on manual inspection and operator judgment, and the data also reported the size of the crack in the material. Do the components classified as unflawed have a smaller mean crack size than components classified as flawed? The results in terms of crack size (in inches) are stored in **Crack**. (Data extracted from B. D. Olin and W. Q. Meeker, "Applications of Statistical Methods to Nondestructive Evaluation," *Technometrics*, 38, 1996, p. 101.)

a. Assuming that the population variances are equal, is there evidence that the mean crack size is smaller for the unflawed specimens than for the flawed specimens? (Use $\alpha = 0.05$.)
b. Repeat (a), assuming that the population variances are not equal.
c. Compare the results of (a) and (b).

10.2 Comparing the Means of Two Related Populations

The hypothesis-testing procedures presented in Section 10.1 enable you to make comparisons and examine differences in the means of two *independent* populations. In this section, you will learn about a procedure for analyzing the difference between the means of two populations when you collect sample data from populations that are related—that is, when results of the first population are *not* independent of the results of the second population.

There are two situations that involve related data between populations. Either you take repeated measurements from the same set of items or individuals or you match items or individuals according to some characteristic. In either situation, you are interested in the *difference between the two related values* rather than the *individual values* themselves.

When you take **repeated measurements** on the same items or individuals, you assume that the same items or individuals will behave alike if treated alike. Your objective is to show that any differences between two measurements of the same items or individuals are due to different treatment conditions. For example, when performing a taste-testing experiment comparing two beverages, you can use each person in the sample as his or her own control so that you can have *repeated measurements* on the same individual.

Another example of repeated measurements involves the pricing of the same goods from two different vendors. For example, have you ever wondered whether new textbook prices at a local college bookstore are different from the prices offered at a major online retailer? You could take two independent samples, that is, select two different sets of textbooks, and then use the hypothesis tests discussed in Section 10.1.

However, by random chance, the first sample may have many large-format hardcover textbooks and the second sample may have many small trade paperback books. This would imply that the first set of textbooks will always be more expensive than the second set of textbooks, regardless of where they are purchased. That observation means that using the Section 10.1 tests would not be a good choice. The better choice would be to use two related samples, that is, to determine the price of the same sample of textbooks at both the local bookstore and the online retailer.

The second situation that involves related data between populations is when you have **matched samples**. Here items or individuals are paired together according to some characteristic of interest. For example, in test marketing a product in two different advertising campaigns, a sample of test markets can be *matched* on the basis of the test market population size and/or demographic variables. By accounting for the differences in test market population size and/or demographic variables, you are better able to measure the effects of the two different advertising campaigns.

Regardless of whether you have matched samples or repeated measurements, the objective is to study the difference between two measurements by reducing the effect of the variability that is due to the items or individuals themselves. Table 10.3 shows the differences in the individual values for two related populations. To read this table, let $X_{11}, X_{12}, \ldots, X_{1n}$ represent the n values from a sample. And let $X_{21}, X_{22}, \ldots, X_{2n}$ represent either the corresponding n matched values from a second sample or the corresponding n repeated measurements from the initial sample. Then, $D_1, D_2, \ldots, D_n$ will represent the corresponding set of n difference *scores* such that

$$D_1 = X_{11} - X_{21}, D_2 = X_{12} - X_{22}, \ldots, \text{and } D_n = X_{1n} - X_{2n}.$$

To test for the mean difference between two related populations, you treat the difference scores, each D_i, as values from a single sample.

TABLE 10.3

Determining the Difference Between Two Related Samples

	Sample		
Value	**1**	**2**	**Difference**
1	X_{11}	X_{21}	$D_1 = X_{11} - X_{21}$
2	X_{12}	X_{22}	$D_2 = X_{12} - X_{22}$
.	.	.	.
.	.	.	.
.	.	.	.
i	X_{1i}	X_{2i}	$D_i = X_{1i} - X_{2i}$
.	.	.	.
.	.	.	.
.	.	.	.
n	X_{1n}	X_{2n}	$D_n = X_{1n} - X_{2n}$

Paired *t* Test

If you assume that the difference scores are randomly and independently selected from a population that is normally distributed, you can use the **paired *t* test for the mean difference** in related populations to determine if there is a significant population mean difference. As with the one-sample t test developed in Section 9.2 [see Equation (9.2) on page 338], the paired t test statistic follows the t distribution with $n - 1$ degrees of freedom. Although the paired t test assumes that the population is normally distributed, you can use this test as long as the sample size is not very small and the population is not highly skewed.

To test the null hypothesis that there is no difference in the means of two related populations:

$$H_0: \mu_D = 0 \text{ (where } \mu_D = \mu_1 - \mu_2)$$

against the alternative that the means are not the same:

$$H_1: \mu_D \neq 0$$

you compute the t_{STAT} test statistic using Equation (10.5).

PAIRED t TEST FOR THE MEAN DIFFERENCE

$$t_{STAT} = \frac{\overline{D} - \mu_D}{\dfrac{S_D}{\sqrt{n}}} \tag{10.5}$$

where

$$\mu_D = \text{hypothesized mean difference}$$

$$\overline{D} = \frac{\sum_{i=1}^{n} D_i}{n}$$

$$S_D = \sqrt{\frac{\sum_{i=1}^{n} (D_i - \overline{D})^2}{n - 1}}$$

The t_{STAT} test statistic follows a t distribution with $n - 1$ degrees of freedom.

For a two-tail test with a given level of significance, α, you reject the null hypothesis if the computed t_{STAT} test statistic is greater than the upper-tail critical value $t_{\alpha/2}$ from the t distribution, or if the computed t_{STAT} test statistic is less than the lower-tail critical value $-t_{\alpha/2}$ from the t distribution. The decision rule is

$$\text{Reject } H_0 \text{ if } t_{STAT} > t_{\alpha/2}$$
$$\text{or if } t_{STAT} < -t_{\alpha/2};$$
$$\text{otherwise, do not reject } H_0.$$

You can use the paired t test for the mean difference to investigate a question raised earlier in this section: Are new textbook prices at a local college bookstore different from the prices offered at a major online retailer?

In this repeated-measurements experiment, you use one set of textbooks. For each textbook, you determine the price at the local bookstore and the price at the online retailer. By determining the two prices for the same textbooks, you can reduce the variability in the prices compared with what would occur if you used two independent sets of textbooks. This approach focuses on the differences between the prices of the same textbooks offered by the two retailers.

You collect data by conducting an experiment from a sample of $n = 19$ textbooks used primarily in business school courses during the summer 2010 semester at a local college. You determine the college bookstore price and the online price (which includes shipping costs, if any). You organize and store the data in **BookPrices**. Table 10.4 shows the results.

TABLE 10.4

Prices of Textbooks at the College Bookstore and at an Online Retailer

Author	Title	Bookstore	Online
Pride	Business 10/e	132.75	136.91
Carroll	Business and Society	201.50	178.58
Quinn	Ethics for the Information Age	80.00	65.00
Bade	Foundations of Microeconomics 5/e	153.50	120.43
Case	Principles of Macroeconomics 9/e	153.50	217.99
Brigham	Financial Management 13/e	216.00	197.10
Griffin	Organizational Behavior 9/e	199.75	168.71
George	Understanding and Managing Organizational Behavior 5/e	147.00	178.63
Grewal	Marketing 2/e	132.00	95.89
Barlow	Abnormal Psychology	182.25	145.49
Foner	Give Me Liberty: Seagull Ed. (V2) 2/e	45.50	37.60
Federer	Mathematical Interest Theory 2/e	89.95	91.69
Hoyle	Advanced Accounting 9/e	123.02	148.41
Haviland	Talking About People 4/e	57.50	53.93
Fuller	Information Systems Project Management	88.25	83.69
Pindyck	Macroeconomics 7/e	189.25	133.32
Mankiw	Macroeconomics 7/e	179.25	151.48
Shapiro	Multinational Financial Management 9/e	210.25	147.30
Losco	American Government 2010 Edition	66.75	55.16

Your objective is to determine whether there is any difference in the mean textbook price between the college bookstore and the online retailer. In other words, is there evidence that the mean price is different between the two sellers of textbooks? Thus, the null and alternative hypotheses are

H_0: $\mu_D = 0$ (There is no difference in the mean price between the college bookstore and the online retailer.)

H_1: $\mu_D \neq 0$ (There is a difference in the mean price between the college bookstore and online retailer.)

Choosing the level of significance $\alpha = 0.05$ and assuming that the differences are normally distributed, you use the paired t test [Equation (10.5)]. For a sample of $n = 19$ textbooks, there are $n - 1 = 18$ degrees of freedom. Using Table E.3, the decision rule is

$$\text{Reject } H_0 \text{ if } t_{STAT} > 2.1009$$
$$\text{or if } t_{STAT} < -2.1009;$$
$$\text{otherwise, do not reject } H_0.$$

For the $n = 19$ differences (see Table 10.4), the sample mean difference is

$$\overline{D} = \frac{\sum_{i=1}^{n} D_i}{n} = \frac{240.66}{19} = 12.6663$$

and

$$S_D = \sqrt{\frac{\sum_{i=1}^{n}(D_i - \overline{D})^2}{n - 1}} = 30.4488$$

From Equation (10.5) on page 378,

$$t_{STAT} = \frac{\overline{D} - \mu_D}{\frac{S_D}{\sqrt{n}}} = \frac{12.6663 - 0}{\frac{30.4488}{\sqrt{19}}} = 1.8132$$

Because $-2.1009 < t_{STAT} = 1.8132 < 2.1009$, you do not reject the null hypothesis, H_0 (see Figure 10.8). There is insufficient evidence of a difference in the mean price of textbooks purchased at the college bookstore and the online retailer.

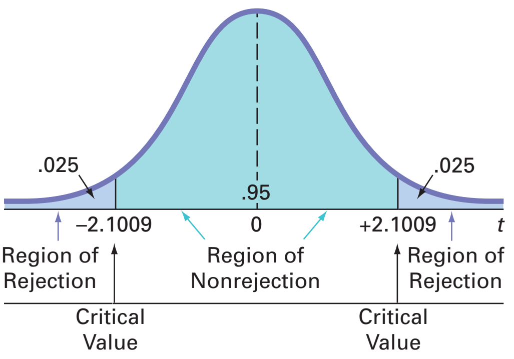


FIGURE 10.8

Two-tail paired t test at the 0.05 level of significance with 18 degrees of freedom

Figure 10.9 presents the results for this example, computing both the t test statistic and the p-value. Because the p-value $= 0.0865 > \alpha = 0.05$, you do not reject H_0. The p-value indicates that if the two sources for textbooks have the same population mean price, the probability that one source would have a sample mean \$12.67 more than the other is 0.0865. Because this probability is greater than $\alpha = 0.05$, you conclude that there is insufficient evidence to reject the null hypothesis.

FIGURE 10.9

Excel and Minitab paired *t* test results for the textbook price data

	A	B	
1	**Paired *t* Test**		
2			
3	**Data**		
4	Hypothesized Mean Diff.	0	
5	Level of significance	0.05	
6			
7	Intermediate Calculations		
8	Sample Size	19	=COUNT(PtCalcs!A:A)
9	DBar	12.6663	=AVERAGE(PtCalcs!C:C)
10	degrees of freedom	18	=B8 - 1
11	S_D	30.4488	=SQRT(SUM(PtCalcs!D:D)/B10)
12	Standard Error	6.9854	=B11/SQRT(A7)
13	*t* Test Statistic	1.8132	=(B9 - B4)/B12
14			
15	**Two-Tail Test**		
16	Lower Critical Value	-2.1009	=-TINV(B5, B10)
17	Upper Critical Value	2.1009	=TINV(B5, B10)
18	*p*-Value	0.0865	=TDIST(ABS(B13), B10, 2)
19	Do not reject the null hypothesis		=IF(B18 < B5, D27, D28)

Not shown
Cell D27: Reject the null hypothesis
Cell D28: Do not reject the null hypothesis

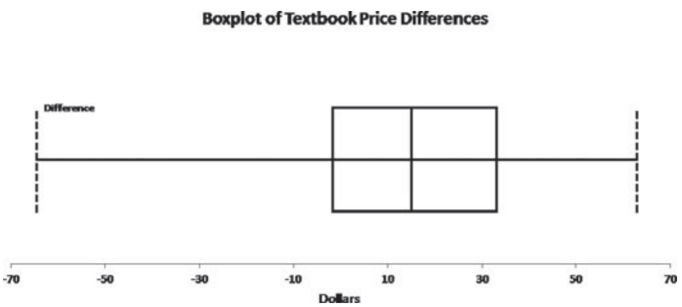
Paired T-Test and CI: Bookstore, Online

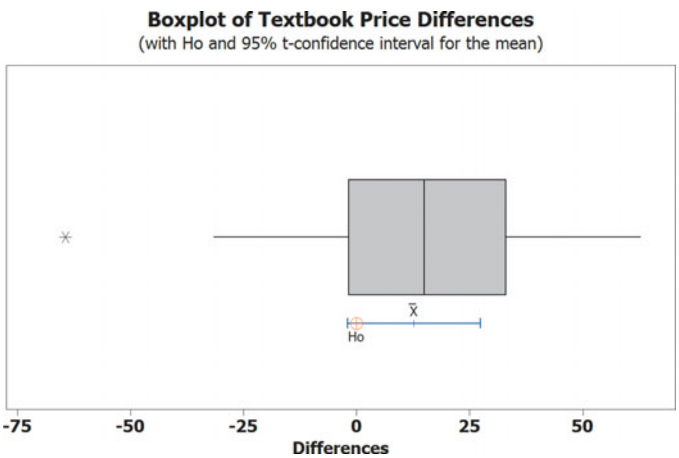
```
Paired T for Bookstore - Online

             N   Mean  StDev  SE Mean
Bookstore   19  139.4   55.0     12.6
Online      19  126.7   52.0     11.9
Difference  19  12.67  30.45     6.99

95% CI for mean difference: (-2.01, 27.34)
T-Test of mean difference = 0 (vs not = 0): T-Value = 1.81  P-Value = 0.087
```

To evaluate the validity of the assumption of normality, you construct a boxplot of the differences, as shown in Figure 10.10.

FIGURE 10.10

Excel and Minitab boxplots for the textbook price data





*For an Excel boxplot of the differences, use column C of the **PtCalcs worksheet**, discussed in the Section EG10.2 In-Depth Excel instructions.*

The Figure 10.10 boxplots show approximate symmetry except for one extreme value. Thus, the data do not greatly contradict the underlying assumption of normality. If a boxplot, histogram, or normal probability plot reveals that the assumption of underlying normality in the population is severely violated, then the *t* test may be inappropriate, especially if the sample size is small. If you believe that the *t* test is inappropriate, you can use either a *nonparametric* procedure that does not make the assumption of underlying normality (see Section 12.8) or make a data transformation (see reference 6) and then recheck the assumptions to determine whether you should use the *t* test.

EXAMPLE 10.2

Paired *t* Test of Pizza Delivery Times

Recall from Example 10.1 on page 369 that a local pizza restaurant situated across the street from your college campus advertises that it delivers to the dormitories faster than the local branch of a national pizza chain. In order to determine whether this advertisement is valid, you and some friends have decided to order 10 pizzas from the local pizza restaurant and 10 pizzas from the national chain. In fact, each time you ordered a pizza from the local pizza restaurant, at the same time, your friends ordered a pizza from the national pizza chain. Thus, you have matched samples. For each of the 10 times that pizzas were ordered, you have one measurement from the local pizza restaurant and one from the national chain. At the 0.05 level of significance, is the mean delivery time for the local pizza restaurant less than the mean delivery time for the national pizza chain?

SOLUTION Use the paired *t* test to analyze the Table 10.5 data (stored in **PizzaTime**). Figure 10.11 shows the paired *t* test results for the pizza delivery data.

TABLE 10.5

Delivery Times for Local Pizza Restaurant and National Pizza Chain

Time	Local	Chain	Difference
1	16.8	22.0	−5.2
2	11.7	15.2	−3.5
3	15.6	18.7	−3.1
4	16.7	15.6	1.1
5	17.5	20.8	−3.3
6	18.1	19.5	−1.4
7	14.1	17.0	−2.9
8	21.8	19.5	2.3
9	13.9	16.5	−2.6
10	20.8	24.0	−3.2
			−21.8

FIGURE 10.11

Excel and Minitab paired *t* test results for the pizza delivery data

	A	B
1	Paired *t* Test for Pizza Delivery Data	
2		
3	Data	
4	Hypothesized Mean Diff.	0
5	Level of significance	0.05
6		
7	Intermediate Calculations	
8	Sample Size	10
9	DBar	-2.1800
10	degrees of freedom	9
11	S_D	2.2641
12	Standard Error	0.7160
13	*t* Test Statistic	-3.0448
14		
15	Lower-Tail Test	
16	Lower Critical Value	-1.8331
17	*p*-Value	0.0070
18	Reject the null hypothesis	

Paired T-Test and CI: Local, Chain

```
Paired T for Local - Chain

              N    Mean  StDev  SE Mean
Local        10  16.700  3.096    0.979
Chain        10  18.880  2.866    0.906
Difference   10  -2.180  2.264    0.716

95% upper bound for mean difference: -0.868
T-Test of mean difference = 0 (vs < 0): T-Value = -3.04  P-Value = 0.007
```

The null and alternative hypotheses are

H_0: $\mu_D \geq 0$ (Mean delivery time for the local pizza restaurant is greater than or equal to the mean delivery time for the national pizza chain.)

H_1: $\mu_D < 0$ (Mean delivery time for the local pizza restaurant is less than the mean delivery time for the national pizza chain.)

Choosing the level of significance $\alpha = 0.05$ and assuming that the differences are normally distributed, you use the paired t test [Equation (10.5) on page 378]. For a sample of $n = 10$ delivery times, there are $n - 1 = 9$ degrees of freedom. Using Table E.3, the decision rule is

$$\text{Reject } H_0 \text{ if } t_{STAT} < -t_{0.05} = -1.8331;$$
$$\text{otherwise, do not reject } H_0.$$

To illustrate the computations, for $n = 10$ differences (see Table 10.5), the sample mean difference is

$$\overline{D} = \frac{\sum_{i=1}^{n} D_i}{n} = \frac{-21.8}{10} = -2.18$$

and the sample standard deviation of the difference is

$$S_D = \sqrt{\frac{\sum_{i=1}^{n}(D_i - \overline{D})^2}{n - 1}} = 2.2641$$

From Equation (10.5) on page 378,

$$t_{STAT} = \frac{\overline{D} - \mu_D}{\frac{S_D}{\sqrt{n}}} = \frac{-2.18 - 0}{\frac{2.2641}{\sqrt{10}}} = -3.0448$$

Because $t_{STAT} = -3.0448$ is less than -1.8331, you reject the null hypothesis, H_0 (the p-value is $0.0070 < 0.05$). There is evidence that the mean delivery time is lower for the local pizza restaurant than for the national pizza chain.

This conclusion is different from the one you reached in Example 10.1 on page 369 when you used the pooled-variance t test for these data. By pairing the delivery times, you are able to focus on the differences between the two pizza delivery services and not the variability created by ordering pizzas at different times of day. The paired t test is a more powerful statistical procedure that is better able to detect the difference between the two pizza delivery services because you are controlling for the time of day they were ordered.

Confidence Interval Estimate for the Mean Difference

Instead of, or in addition to, testing for the difference between the means of two related populations, you can use Equation (10.6) to construct a confidence interval estimate for the mean difference.

CONFIDENCE INTERVAL ESTIMATE FOR THE MEAN DIFFERENCE

$$\overline{D} \pm t_{\alpha/2}\frac{S_D}{\sqrt{n}}$$

or

$$\overline{D} - t_{\alpha/2}\frac{S_D}{\sqrt{n}} \leq \mu_D \leq \overline{D} + t_{\alpha/2}\frac{S_D}{\sqrt{n}} \quad \textbf{(10.6)}$$

where $t_{\alpha/2}$ is the critical value of the t distribution, with $n - 1$ degrees of freedom, for an area of $\alpha/2$ in the upper tail.

Recall the example comparing textbook prices on page 379. Using Equation (10.6), $\overline{D} = 12.6663, S_D = 30.4488, n = 19$, and $t_{\alpha/2} = 2.1009$ (for 95% confidence and $n - 1 = 18$ degrees of freedom),

$$12.6663 \pm (2.1009)\frac{30.4488}{\sqrt{19}}$$

$$12.6663 \pm 14.6757$$

$$-2.0094 \leq \mu_D \leq 27.342$$

Thus, with 95% confidence, the mean difference in textbook prices between the college bookstore and the online retailer is between –\$2.0094 and \$27.342. Because the interval estimate contains zero, you can conclude that there is insufficient evidence of a difference in the population means. There is insufficient evidence of a difference in the mean prices of textbooks at the college bookstore and the online retailer.

Problems for Section 10.2

LEARNING THE BASICS

10.18 An experimental design for a paired t test has 20 pairs of identical twins. How many degrees of freedom are there in this t test?

10.19 Fifteen volunteers are recruited to participate in an experiment. A measurement is made (such as blood pressure) before each volunteer is asked to read a particularly upsetting passage from a book and after each volunteer reads the passage from the book. In the analysis of the data collected from this experiment, how many degrees of freedom are there in the test?

APPLYING THE CONCEPTS

SELF Test **10.20** Nine experts rated two brands of Colombian coffee in a taste-testing experiment. A rating on a 7-point scale (1 = extremely unpleasing, 7 = extremely pleasing) is given for each of four characteristics: taste, aroma, richness, and acidity. The following data (stored in **Coffee**) display the ratings accumulated over all four characteristics.

	Brand	
Expert	**A**	**B**
C.C.	24	26
S.E.	27	27
E.G.	19	22
B.L.	24	27
C.M.	22	25
C.N.	26	27
G.N.	27	26
R.M.	25	27
P.V.	22	23

a. At the 0.05 level of significance, is there evidence of a difference in the mean ratings between the two brands?
b. What assumption is necessary about the population distribution in order to perform this test?
c. Determine the p-value in (a) and interpret its meaning.
d. Construct and interpret a 95% confidence interval estimate of the difference in the mean ratings between the two brands.

10.21 In industrial settings, alternative methods often exist for measuring variables of interest. The data in **Measurement** (coded to maintain confidentiality) represent measurements in-line that were collected from an analyzer during the production process and from an analytical lab. (Data extracted from M. Leitnaker, "Comparing Measurement Processes: In-line Versus Analytical Measurements," *Quality Engineering*, 13, 2000–2001, pp. 293–298.)
a. At the 0.05 level of significance, is there evidence of a difference in the mean measurements in-line and from an analytical lab?
b. What assumption is necessary about the population distribution in order to perform this test?
c. Use a graphical method to evaluate the validity of the assumption in (a).
d. Construct and interpret a 95% confidence interval estimate of the difference in the mean measurements in-line and from an analytical lab.

10.22 Is there a difference in the prices at a warehouse club such as Costco and store brands? To investigate this, a random sample of 10 purchases was selected, and the prices were compared. (Data extracted from "Shop Smart and Save Big," *Consumer Reports*, May 2009, p. 17.) The prices for the products are stored in **Shopping1**.
a. At the 0.05 level of significance, is there evidence of a difference between the mean price of Costco purchases and store-brand purchases?
b. What assumption is necessary about the population distribution in order to perform this test?

c. Construct a 95% confidence interval estimate of the mean difference in price between Costco and store brands. Interpret the interval.
d. Compare the results of (a) and (c).

10.23 In tough economic times, the business staff at magazines are challenged to sell advertising space in their publications. Thus, one indicator of a weak economy is the decline in the number of "ad pages" that magazines have sold. The file **Ad Pages** contains the number of ad pages found in the May 2008 and May 2009 issues of 12 men's magazines. (Data extracted from W. Levith, "Magazine Monitor," *Mediaweek*, April 20, 2009, p. 53.)
a. At the 0.05 level of significance, is there evidence that the mean number of ad pages was higher in May 2008 than in May 2009?
b. What assumption is necessary about the population distribution in order to perform this test?
c. Use a graphical method to evaluate the validity of the assumption in (b).
d. Construct and interpret a 95% confidence interval estimate of the difference in the mean number of ad pages in men's magazines between May 2008 and May 2009.

10.24 Multiple myeloma, or blood plasma cancer, is characterized by increased blood vessel formulation (angiogenesis) in the bone marrow that is a predictive factor in survival. One treatment approach used for multiple myeloma is stem cell transplantation with the patient's own stem cells. The following data (stored in **Myeloma**) represent the bone marrow microvessel density for patients who had a complete response to the stem cell transplant (as measured by blood and urine tests). The measurements were taken immediately prior to the stem cell transplant and at the time the complete response was determined.
a. At the 0.05 level of significance, is there evidence that the mean bone marrow microvessel density is higher before the stem cell transplant than after the stem cell transplant?
b. Interpret the meaning of the p-value in (a).
c. Construct and interpret a 95% confidence interval estimate of the mean difference in bone marrow microvessel density before and after the stem cell transplant.
d. What assumption is necessary about the population distribution in order to perform the test in (a)?

Patient	Before	After
1	158	284
2	189	214
3	202	101
4	353	227
5	416	290
6	426	176
7	441	290

Source: Data extracted from S. V. Rajkumar, R. Fonseca, T. E. Witzig, M. A. Gertz, and P. R. Greipp, "Bone Marrow Angiogenesis in Patients Achieving Complete Response After Stem Cell Transplantation for Multiple Myeloma," *Leukemia,* 1999, 13, pp. 469–472.

10.25 Over the past year, the vice president for human resources at a large medical center has run a series of three-month workshops aimed at increasing worker motivation and performance. To check the effectiveness of the workshops, she selected a random sample of 35 employees from the personnel files. She collected the employee performance ratings recorded before and after workshop attendance and stored the paired ratings, along with descriptive statistics and the results of a paired t test in the **Perform Excel workbook** (**Perform.xls**) and in the **Perform Minitab project** (**Perform.mpj**). Review her results and state your findings and conclusions in a report to the vice president for human resources.

10.26 The data in **Concrete1** represent the compressive strength, in thousands of pounds per square inch (psi), of 40 samples of concrete taken two and seven days after pouring.

Source: Data extracted from O. Carrillo-Gamboa and R. F. Gunst, "Measurement-Error-Model Collinearities," *Technometrics,* 34, 1992, pp. 454–464.

a. At the 0.01 level of significance, is there evidence that the mean strength is lower at two days than at seven days?
b. What assumption is necessary about the population distribution in order to perform this test?
c. Find the p-value in (a) and interpret its meaning.

10.3 Comparing the Proportions of Two Independent Populations

Often, you need to make comparisons and analyze differences between two population proportions. You can perform a test for the difference between two proportions selected from independent populations by using two different methods. This section presents a procedure whose test statistic, Z_{STAT}, is approximated by a standardized normal distribution. In Section 12.1, a procedure whose test statistic, χ^2_{STAT}, is approximated by a chi-square distribution is used. As you will see when you read that section, the results from these two tests are equivalent.

Z Test for the Difference Between Two Proportions

In evaluating differences between two population proportions, you can use a ***Z* test for the difference between two proportions**. The Z_{STAT} test statistic is based on the difference between two sample proportions $(p_1 - p_2)$. This test statistic, given in Equation (10.7), approximately follows a standardized normal distribution for large enough sample sizes.

Z TEST FOR THE DIFFERENCE BETWEEN TWO PROPORTIONS

$$Z_{STAT} = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\bar{p}(1 - \bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \tag{10.7}$$

with

$$\bar{p} = \frac{X_1 + X_2}{n_1 + n_2} \quad p_1 = \frac{X_1}{n_1} \quad p_2 = \frac{X_2}{n_2}$$

where

p_1 = proportion of items of interest in sample 1

X_1 = number of items of interest in sample 1

n_1 = sample size of sample 1

π_1 = proportion of items of interest in population 1

p_2 = proportion of items of interest in sample 2

X_2 = number of items of interest in sample 2

n_2 = sample size of sample 2

π_2 = proportion of items of interest in population 2

$\bar{p}$ = pooled estimate of the population proportion of items of interest

The Z_{STAT} test statistic approximately follows a standardized normal distribution.

Under the null hypothesis in the *Z* test for the difference between two proportions, you assume that the two population proportions are equal $(\pi_1 = \pi_2)$. Because the pooled estimate for the population proportion is based on the null hypothesis, you combine, or pool, the two sample proportions to compute $\bar{p}$, an overall estimate of the common population proportion. This estimate is equal to the number of items of interest in the two samples combined $(X_1 + X_2)$ divided by the total sample size from the two samples combined $(n_1 + n_2)$.

As shown in the following table, you can use this *Z* test for the difference between population proportions to determine whether there is a difference in the proportion of items of interest in the two populations (two-tail test) or whether one population has a higher proportion of items of interest than the other population (one-tail test):

Two-Tail Test	One-Tail Test	One-Tail Test
$H_0: \pi_1 = \pi_2$	$H_0: \pi_1 \geq \pi_2$	$H_0: \pi_1 \leq \pi_2$
$H_1: \pi_1 \neq \pi_2$	$H_1: \pi_1 < \pi_2$	$H_1: \pi_1 > \pi_2$

where

π_1 = proportion of items of interest in population 1

π_2 = proportion of items of interest in population 2

To test the null hypothesis that there is no difference between the proportions of two independent populations:

$$H_0: \pi_1 = \pi_2$$

against the alternative that the two population proportions are not the same:

$$H_1: \pi_1 \neq \pi_2$$

You use the Z_{STAT} test statistic, given by Equation (10.7). For a given level of significance, α, you reject the null hypothesis if the computed Z_{STAT} test statistic is greater than the upper-tail critical value from the standardized normal distribution or if the computed Z_{STAT} test statistic is less than the lower-tail critical value from the standardized normal distribution.

To illustrate the use of the Z test for the equality of two proportions, suppose that you are the manager of T.C. Resort Properties, a collection of five upscale resort hotels located on two tropical islands. On one of the islands, T.C. Resort Properties has two hotels, the Beachcomber and the Windsurfer. You have defined the business objective as improving the return rate of guests at the Beachcomber and the Windsurfer hotels. On the questionnaire completed by hotel guests upon their departure, one question asked is whether the guest is likely to return to the hotel. Responses to this and other questions were collected from 227 guests at the Beachcomber and 262 guests at the Windsurfer. The results for this question indicated that 163 of 227 guests at the Beachcomber responded yes, they were likely to return to the hotel and 154 of 262 guests at the Windsurfer responded yes, they were likely to return to the hotel. At the 0.05 level of significance, is there evidence of a significant difference in guest satisfaction (as measured by the likelihood to return to the hotel) between the two hotels?

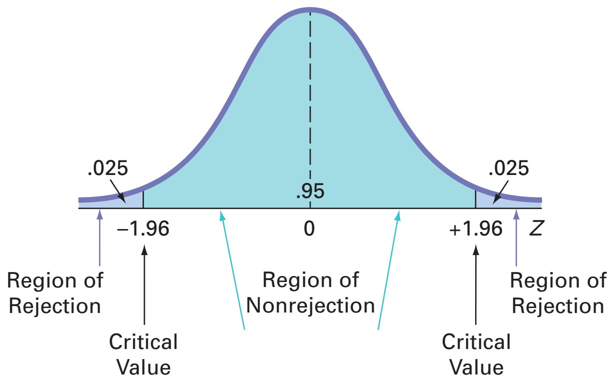
The null and alternative hypotheses are

$$H_0: \pi_1 = \pi_2 \text{ or } \quad \pi_1 - \pi_2 = 0$$
$$H_1: \pi_1 \neq \pi_2 \text{ or } \quad \pi_1 - \pi_2 \neq 0$$

Using the 0.05 level of significance, the critical values are -1.96 and $+1.96$ (see Figure 10.12), and the decision rule is

$$\text{Reject } H_0 \text{ if } Z_{STAT} < -1.96$$
$$\text{or if } Z_{STAT} > +1.96;$$
$$\text{otherwise, do not reject } H_0.$$

FIGURE 10.12
Regions of rejection and nonrejection when testing a hypothesis for the difference between two proportions at the 0.05 level of significance



Using Equation (10.7) on page 386,

$$Z_{STAT} = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\bar{p}(1 - \bar{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

where

$$p_1 = \frac{X_1}{n_1} = \frac{163}{227} = 0.7181 \quad p_2 = \frac{X_2}{n_2} = \frac{154}{262} = 0.5878$$

and

$$\bar{p} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{163 + 154}{227 + 262} = \frac{317}{489} = 0.6483$$

so that

$$Z_{STAT} = \frac{(0.7181 - 0.5878) - (0)}{\sqrt{0.6483(1 - 0.6483)\left(\frac{1}{227} + \frac{1}{262}\right)}}$$

$$= \frac{0.1303}{\sqrt{(0.228)(0.0082)}}$$

$$= \frac{0.1303}{\sqrt{0.00187}}$$

$$= \frac{0.1303}{0.0432} = +3.0088$$

Using the 0.05 level of significance, you reject the null hypothesis because $Z_{STAT} = +3.0088 > +1.96$. The p-value is 0.0026 (computed using Table E.2 or from Figure 10.13) and indicates that if the null hypothesis is true, the probability that a Z_{STAT} test statistic is less than -3.0088 is 0.0013, and, similarly, the probability that a Z_{STAT} test statistic is greater than $+3.0088$ is 0.0013. Thus, for this two-tail test, the p-value is $0.0013 + 0.0013 = 0.0026$. Because $0.0026 < \alpha = 0.05$, you reject the null hypothesis. There is evidence to conclude that the two hotels are significantly different with respect to guest satisfaction; a greater proportion of guests are willing to return to the Beachcomber than to the Windsurfer.

FIGURE 10.13

Excel and Minitab *Z* test results for the difference between two proportions for the hotel guest satisfaction problem

	A	B	
1	Z Test for Differences in Two Proportions		
2			
3	Data		
4	Hypothesized Difference	0	
5	Level of Significance	0.05	
6	Group 1		
7	Number of Items of Interest	163	
8	Sample Size	227	
9	Group 2		
10	Number of Items of Interest	154	
11	Sample Size	262	
12			
13	Intermediate Calculations		
14	Group 1 Proportion	0.7181	=B7/B8
15	Group 2 Proportion	0.5878	=B10/B11
16	Difference in Two Proportions	0.1303	=B14 - B15
17	Average Proportion	0.6483	=(B7 + B10)/(B8 + B11)
18	Z Test Statistic	3.0088	=(B16 - B4)/SQRT(B17 * (1 - B17) * (1/B8 + 1/B11))
19			
20	Two-Tail Test		
21	Lower Critical Value	-1.9600	=NORMSINV(B5/2)
22	Upper Critical Value	1.9600	=NORMSINV(1 - B5/2)
23	p-Value	0.0026	=2 * (1 - NORMSDIST(ABS(B18)))
24	Reject the null hypothesis		=IF(B23 < B5, "Reject the null hypothesis", "Do not reject the null hypothesis")

Test and CI for Two Proportions

```
Sample    X    N  Sample p
1       163  227  0.718062
2       154  262  0.587786

Difference = p (1) - p (2)
Estimate for difference:  0.130275
95% CI for difference:  (0.0467379, 0.213813)
Test for difference = 0 (vs not = 0):  Z = 3.01  P-Value = 0.003

Fisher's exact test: P-Value = 0.003
```

EXAMPLE 10.3

Testing for the Difference Between Two Proportions

A growing concern about privacy on the Internet has led more people to monitor their online identities. (Data extracted from Drilling Down, "Managing Reputations on Social Sites," *The New York Times*, June 14, 2010, pp. B2B.) The survey reported that 44% of Internet users ages 18 to 29 have taken steps to restrict the amount of information available about themselves online as compared to 20% of Internet users older than 65 who have done the same thing. The sample size in each group was not reported. Suppose that the survey consisted of 100 individuals in each age group. At the 0.05 level of significance, is the proportion of Internet users ages 18 to 29 who have taken steps to restrict the amount of information available about themselves online greater than the proportion of Internet users older than 65 who have done the same thing?

SOLUTION Because you want to know whether there is evidence that the proportion in the 18-to-29 age group is *greater* than in the over-65 age group, you have a one-tail test. The null and alternative hypotheses are

H_0: $\pi_1 \le \pi_2$ (The proportion of Internet users ages 18 to 29 who have taken steps to restrict the amount of information available about themselves online is less than or equal to the proportion of Internet users older than 65 who have done the same thing.)

H_1: $\pi_1 > \pi_2$ (The proportion of Internet users ages 18 to 29 who have taken steps to restrict the amount of information available about themselves online is greater than the proportion of Internet users older than 65 who have done the same thing.)

Using the 0.05 level of significance, for the one-tail test in the upper tail, the critical value is +1.645. The decision rule is

Reject H_0 if $Z_{STAT} > +1.645$;

otherwise, do not reject H_0.

Using Equation (10.7) on page 386,

$$Z_{STAT} = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\bar{p}(1 - \bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where

$$p_1 = \frac{X_1}{n_1} = \frac{44}{100} = 0.44 \quad p_2 = \frac{X_2}{n_2} = \frac{20}{100} = 0.20$$

and

$$\bar{p} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{44 + 20}{100 + 100} = \frac{64}{200} = 0.32$$

so that

$$Z_{STAT} = \frac{(0.44 - 0.20) - (0)}{\sqrt{0.32(1 - 0.32)\left(\frac{1}{100} + \frac{1}{100}\right)}}$$

$$= \frac{0.24}{\sqrt{(0.2176)(0.02)}}$$

$$= \frac{0.24}{\sqrt{0.004352}}$$

$$= \frac{0.24}{0.06597} = +3.638$$

Using the 0.05 level of significance, you reject the null hypothesis because $Z_{STAT} = +3.638 > +1.645$. The p-value is approximately 0.0001. Therefore, if the null hypothesis is true, the probability that a Z_{STAT} test statistic is greater than +3.638 is approximately 0.0001 (which is less than $\alpha = 0.05$). You conclude that there is evidence that the proportion of Internet users ages 18 to 29 who have taken steps to restrict the amount of information available about themselves online is greater than the proportion of Internet users older than 65 who have done the same thing.

Confidence Interval Estimate for the Difference Between Two Proportions

Instead of, or in addition to, testing for the difference between the proportions of two independent populations, you can construct a confidence interval estimate for the difference between the two proportions using Equation (10.8).

CONFIDENCE INTERVAL ESTIMATE FOR THE DIFFERENCE BETWEEN TWO PROPORTIONS

$$(p_1 - p_2) \pm Z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

or

$$(p_1 - p_2) - Z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \le (\pi_1 - \pi_2) \le (p_1 - p_2) + Z_{\alpha/2}\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}} \quad \textbf{(10.8)}$$

To construct a 95% confidence interval estimate for the population difference between the proportion of guests who would return to the Beachcomber and who would return to the Windsurfer, you use the results on page 387 or from Figure 10.13 on page 388:

$$p_1 = \frac{X_1}{n_1} = \frac{163}{227} = 0.7181 \qquad p_2 = \frac{X_2}{n_2} = \frac{154}{262} = 0.5878$$

Using Equation (10.8),

$$(0.7181 - 0.5878) \pm (1.96)\sqrt{\frac{0.7181(1-0.7181)}{227} + \frac{0.5878(1-0.5878)}{262}}$$

$$0.1303 \pm (1.96)(0.0426)$$

$$0.1303 \pm 0.0835$$

$$0.0468 \le (\pi_1 - \pi_2) \le 0.2138$$

Thus, you have 95% confidence that the difference between the population proportion of guests who would return to the Beachcomber and the Windsurfer is between 0.0468 and 0.2138. In percentages, the difference is between 4.68% and 21.38%. Guest satisfaction is higher at the Beachcomber than at the Windsurfer.

Problems for Section 10.3

LEARNING THE BASICS

10.27 Let $n_1 = 100, X_1 = 50, n_2 = 100$, and $X_2 = 30$.

a. At the 0.05 level of significance, is there evidence of a significant difference between the two population proportions?

b. Construct a 95% confidence interval estimate for the difference between the two population proportions.

10.28 Let $n_1 = 100, X_1 = 45, n_2 = 50$, and $X_2 = 25$.

a. At the 0.01 level of significance, is there evidence of a significant difference between the two population proportions?

b. Construct a 99% confidence interval estimate for the difference between the two population proportions.

APPLYING THE CONCEPTS

10.29 A survey of 500 shoppers was taken in a large metropolitan area to determine various information about consumer behavior. Among the questions asked was, "Do you enjoy shopping for clothing?" Of 240 males, 136 answered yes. Of 260 females, 224 answered yes.

a. Is there evidence of a significant difference between males and females in the proportion who enjoy shopping for clothing at the 0.01 level of significance?

b. Find the p-value in (a) and interpret its meaning.

c. Construct and interpret a 99% confidence interval estimate for the difference between the proportion of males and females who enjoy shopping for clothing.

d. What are your answers to (a) through (c) if 206 males enjoyed shopping for clothing?

10.30 Does it take more effort to be removed from an email list than it used to? A study of 100 large online retailers revealed the following:

	NEED THREE OR MORE CLICKS TO BE REMOVED	
YEAR	Yes	No
2009	39	61
2008	7	93

Source: Data extracted from "More Clicks to Escape an Email List," *The New York Times,* March 29, 2010, p. B2.

a. Set up the null and alternative hypotheses to try to determine whether it takes more effort to be removed from an email list than it used to.

b. Conduct the hypothesis test defined in (a), using the 0.05 level of significance.

c. Does the result of your test in (b) make it appropriate to claim that it takes more effort to be removed from an email list than it used to?

10.31 Some people enjoy the *anticipation* of an upcoming product or event and prefer to pay in advance and delay the actual consumption/delivery date. In other cases, people do not want a delay. An article in the *Journal of Marketing Research* reported on an experiment in which 50 individuals were told that they had just purchased a ticket to a concert and 50 were told that they had just purchased a personal digital assistant (PDA). The participants were then asked to indicate their preferences for attending the concert or receiving the PDA. Did they prefer tonight or tomorrow, or would they prefer to wait two to four weeks? The individuals were told to ignore their schedule constraints in order to better measure their willingness to delay the consumption/delivery of their purchase. The following table gives partial results of the study:

When to Receive Purchase	Concert	PDA
Tonight or tomorrow	28	47
Two to four weeks	22	3
Total	50	50

Source: Data adapted from O. Amir and D. Ariely, "Decisions by Rules: The Case of Unwillingness to Pay for Beneficial Delays," *Journal of Marketing Research,* February 2007, Vol. XLIV, pp. 142–152.

a. What proportion of the participants would prefer to delay the date of the concert?

b. What proportion of the participants would prefer to delay receipt of a new PDA?

c. Using the 0.05 level of significance, is there evidence of a significant difference in the proportion willing to delay the date of the concert and the proportion willing to delay receipt of a new PDA?

SELF Test **10.32** Do people of different age groups differ in their beliefs about response time to email messages? A survey by the Center for the Digital Future of the University of Southern California reported that 70.7% of users over 70 years of age believe that email messages should be answered quickly as compared to 53.6% of users 12 to 50 years old. (Data extracted from A. Mindlin, "Older E-mail Users Favor Fast Replies," *The New York Times*, July 14, 2008, p. B3.) Suppose that the survey was based on 1,000 users over 70 years of age and 1,000 users 12 to 50 years old.

a. At the 0.01 level of significance, is there evidence of a significant difference between the two age groups in the proportion that believe that email messages should be answered quickly?

b. Find the p-value in (a) and interpret its meaning.

10.33 A survey was conducted of 665 consumer magazines on the practices of their websites. Of these, 273 magazines reported that online-only content is copy-edited as rigorously as print content; 379 reported that online-only content is fact-checked as rigorously as print content. (Data extracted from S. Clifford, "Columbia Survey Finds a Slack

Editing Process of Magazine Web Sites," *The New York Times*, March 1, 2010, p. B6.) Suppose that a sample of 500 newspapers revealed that 252 reported that online-only content is copy-edited as rigorously as print content and 296 reported that online-only content is fact-checked as rigorously as print content.

a. At the 0.05 level of significance, is there evidence of a difference between consumer magazines and newspapers in the proportion of online-only content that is copy-edited as rigorously as print content ?

b. Find the p-value in (a) and interpret its meaning.

c. At the 0.05 level of significance, is there evidence of a difference between consumer magazines and newspapers in the proportion of online-only content that is fact-checked as rigorously as print content?

10.34 How do Americans feel about ads on websites? A survey of 1,000 adult Internet users found that 670 opposed ads on websites. (Data extracted from S. Clifford, "Tacked for Ads? Many Americans Say No Thanks," *The New York Times*, September 30, 2009, p. B3.). Suppose that a survey of 1,000 Internet users age 12–17 found that 510 opposed ads on websites

a. At the 0.05 level of significance, is there evidence of a difference between adult Internet users and Internet users age 12–17 in the proportion who oppose ads?

b. Find the p-value in (a) and interpret its meaning.

10.35 Where people turn for news is different for various age groups. (Data extracted from "Cellphone Users Who Access News on Their Phones," *USA Today*, March 1, 2010, p. 1A.) A study was conducted on the use of cell phones for accessing news. The study reported that 47% of users under age 50 and 15% of users age 50 and over accessed news on their cell phones. Suppose that the survey consisted of 1,000 users under age 50, of whom 470 accessed news on their cell phones, and 891 users age 50 and over, of whom 134 accessed news on their cell phones.

a. Is there evidence of a significant difference in the proportion of users under age 50 and users 50 years and older that accessed the news on their cell phones? (Use $\alpha = 0.05$.)

b. Determine the p-value in (a) and interpret its meaning.

c. Construct and interpret a 95% confidence interval estimate for the difference between the population proportion of users under 50 years old and those 50 years or older who access the news on their cell phones.

10.4 *F* Test for the Ratio of Two Variances

Often you need to determine whether two independent populations have the same variability. By testing variances, you can detect differences in the variability in two independent populations. One important reason to test for the difference between the variances of two populations is to determine whether to use the pooled-variance t test (which assumes equal variances) or the separate-variance t test (which does not assume equal variances) while comparing the means of two independent populations.

The test for the difference between the variances of two independent populations is based on the ratio of the two sample variances. If you assume that each population is normally distributed, then the ratio S_1^2/S_2^2 follows the F distribution (see Table E.5). The critical values of the ***F* distribution** in Table E.5 depend on the degrees of freedom in the two samples. The degrees of freedom in the numerator of the ratio are for the first sample, and the degrees of freedom in the denominator are for the second sample. The first sample taken from the first population is defined as the sample that has the *larger* sample variance. The second sample taken from the second population is the sample with the *smaller* sample variance. Equation (10.9) defines the ***F* test for the ratio of two variances**.

F TEST STATISTIC FOR TESTING THE RATIO OF TWO VARIANCES

The F_{STAT} test statistic is equal to the variance of sample 1 (the larger sample variance) divided by the variance of sample 2 (the smaller sample variance).

$$F_{STAT} = \frac{S_1^2}{S_2^2} \quad \textbf{(10.9)}$$

where

S_1^2 = variance of sample 1 (the larger sample variance)

S_2^2 = variance of sample 2 (the smaller sample variance)

$$n_1 = \text{sample size selected from population 1}$$
$$n_2 = \text{sample size selected from population 2}$$
$$n_1 - 1 = \text{degrees of freedom from sample 1 (i.e., the numerator degrees of freedom)}$$
$$n_2 - 1 = \text{degrees of freedom from sample 2 (i.e., the denominator degrees of freedom)}$$

The F_{STAT} test statistic follows an F distribution with $n_1 - 1$ and $n_2 - 1$ degrees of freedom.

For a given level of significance, α, to test the null hypothesis of equality of population variances:

$$H_0: \sigma_1^2 = \sigma_2^2$$

against the alternative hypothesis that the two population variances are not equal:

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

you reject the null hypothesis if the computed F_{STAT} test statistic is greater than the upper-tail critical value, $F_{\alpha/2}$, from the F distribution, with $n_1 - 1$ degrees of freedom in the numerator and $n_2 - 1$ degrees of freedom in the denominator. Thus, the decision rule is

$$\text{Reject } H_0 \text{ if } F_{STAT} > F_{\alpha/2};$$
$$\text{otherwise, do not reject } H_0.$$

To illustrate how to use the F test to determine whether the two variances are equal, return to the BLK Beverages scenario on page 365 concerning the sales of BLK Cola in two different display locations. To determine whether to use the pooled-variance t test or the separate-variance t test in Section 10.1, you can test the equality of the two population variances. The null and alternative hypotheses are

$$H_0: \sigma_1^2 = \sigma_2^2$$
$$H_1: \sigma_1^2 \neq \sigma_2^2$$

Because you are defining sample 1 as the group with the larger sample variance, the rejection region in the upper tail of the F distribution contains $\alpha/2$. Using the level of significance $\alpha = 0.05$, the rejection region in the upper tail contains 0.025 of the distribution.

Because there are samples of 10 stores for each of the two display locations, there are $10-1 = 9$ degrees of freedom in the numerator (the sample with the larger variance) and also in the denominator (the sample with the smaller variance). $F_{\alpha/2}$, the upper-tail critical value of the F distribution, is found directly from Table E.5, a portion of which is presented in Table 10.6. Because there are 9 degrees of freedom in the numerator and 9 degrees of freedom in the denominator, you find the upper-tail critical value, $F_{\alpha/2}$, by looking in the column labeled 9 and the row labeled 9. Thus, the upper-tail critical value of this F distribution is 4.03. Therefore, the decision rule is

$$\text{Reject } H_0 \text{ if } F_{STAT} > F_{0.025} = 4.03;$$
$$\text{otherwise, do not reject } H_0.$$

TABLE 10.6

Finding the Upper-Tail Critical Value of F with 9 and 9 Degrees of Freedom for an Upper-Tail Area of 0.025

	Cumulative Probabilities = 0.975						
	Upper-Tail Area = 0.025						
	Numerator df_1						
Denominator df_2	**1**	**2**	**3**	**...**	**7**	**8**	$\boxed{\mathbf{9}}$
1	647.80	799.50	864.20	...	948.20	956.70	963.30
2	38.51	39.00	39.17	...	39.36	39.37	39.39
3	17.44	16.04	15.44	...	14.62	14.54	14.47
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.
7	8.07	6.54	5.89	...	4.99	4.90	4.82
8	7.57	6.06	5.42	...	4.53	4.43	4.36
$\boxed{\mathbf{9}}$	7.21	5.71	5.08	...	4.20	4.10	$\boxed{4.03}$

Source: Extracted from Table E.5.

Using Equation (10.9) on page 392 and the cola sales data (see Table 10.1 on page 367),

$$S_1^2 = (18.7264)^2 = 350.6778 \quad S_2^2 = (12.5433)^2 = 157.3333$$

so that

$$F_{STAT} = \frac{S_1^2}{S_2^2}$$

$$= \frac{350.6778}{157.3333} = 2.2289$$

Because $F_{STAT} = 2.2289 < 4.03$, you do not reject H_0. Figure 10.14 shows the results for this test, including the p-value, 0.248. Because $0.248 > 0.05$, you conclude that there is no evidence of a significant difference in the variability of the sales of cola for the two display locations.

FIGURE 10.14

Excel and Minitab F test results for the BLK Cola sales data

	A	B	
1	F Test for Differences in Two Variances		
2			
3	Data		
4	Level of Significance	0.05	
5	Larger-Variance Sample		
6	Sample Size	10	=COUNT(DATACOPY!$A:$A)
7	Sample Variance	350.6778	=VAR(DATACOPY!$A:$A)
8	Smaller-Variance Sample		
9	Sample Size	10	=COUNT(DATACOPY!$B:$B)
10	Sample Variance	157.3333	=VAR(DATACOPY!$B:$B)
11			
12	Intermediate Calculations		
13	*F* Test Statistic	2.2289	=B7/B10
14	Population 1 Sample Degrees of Freedom	9	=B6 - 1
15	Population 2 Sample Degrees of Freedom	9	=B9 - 1
16			
17	Two-Tail Test		
18	Upper Critical Value	4.0260	=FINV(B4/2, B14, B15)
19	*p*-Value	0.2482	=2 * E17
20	Do not reject the null hypothesis		=IF(B19 < B4, "Reject the null hypothesis", "Do not reject the null hypothesis")

Not shown
Cell E17: =FDIST(B13, B14, B15)

Test and CI for Two Variances: Normal, End-Aisle

```
Method
Null hypothesis         Sigma(Normal) / Sigma(End-Aisle) = 1
Alternative hypothesis  Sigma(Normal) / Sigma(End-Aisle) not = 1
Significance level      Alpha = 0.05


Statistics
Variable    N   StDev  Variance
Normal     10  18.726   350.678
End-Aisle  10  12.543   157.333

Ratio of standard deviations = 1.493
Ratio of variances = 2.229


95% Confidence Intervals
                                   CI for
Distribution    CI for StDev      Variance
of Data            Ratio           Ratio
Normal         (0.744, 2.996)  (0.554, 8.973)
Continuous     (0.664, 3.082)  (0.441, 9.497)


Tests
                                                  Test
Method                          DF1  DF2  Statistic  P-Value
F Test (normal)                   9    9       2.23    0.248
Levene's Test (any continuous)    1   18       1.27    0.275
```

In testing for a difference between two variances using the F test described in this section, you assume that each of the two populations is normally distributed. The F test is very sensitive to the normality assumption. If boxplots or normal probability plots suggest even a mild departure from normality for either of the two populations, you should not use the F test.

If this happens, you should use the Levene test (see Section 11.1) or a nonparametric approach (see references 1 and 2).

In testing for the equality of variances as part of assessing the validity of the pooled-variance t test procedure, the F test is a two-tail test with $\alpha/2$ in the upper tail. However, when you are interested in examining the variability in situations other than the pooled-variance t test, the F test is often a one-tail test. Example 10.4 illustrates a one-tail test.

EXAMPLE 10.4

A One-Tail Test for the Difference Between Two Variances

A professor in the accounting department of a business school would like to determine whether there is more variability in the final exam scores of students taking the introductory accounting course who are not majoring in accounting than for students taking the course who are majoring in accounting. Random samples of 13 non-accounting majors and 10 accounting majors are selected from the professor's class roster in his large lecture, and the following results are computed based on the final exam scores:

$$\text{Non-accounting}: \quad n_1 = 13 \quad S_1^2 = 210.2$$
$$\text{Accounting}: \quad n_2 = 10 \quad S_2^2 = 36.5$$

At the 0.05 level of significance, is there evidence that there is more variability in the final exam scores of students taking the introductory accounting course who are not majoring in accounting than for students taking the course who are majoring in accounting? Assume that the population final exam scores are normally distributed.

SOLUTION The null and alternative hypotheses are

$$H_0: \sigma_{NA}^2 \leq \sigma_A^2$$
$$H_1: \sigma_{NA}^2 > \sigma_A^2$$

The F_{STAT} test statistic is given by Equation (10.9) on page 392:

$$F_{STAT} = \frac{S_1^2}{S_2^2}$$

You use Table E.5 to find the upper critical value of the F distribution. With $n_1 - 1 = 13 - 1 = 12$ degrees of freedom in the numerator, $n_2 - 1 = 10 - 1 = 9$ degrees of freedom in the denominator, and $\alpha = 0.05$, the upper-tail critical value, $F_{0.05}$, is 3.07. The decision rule is

$$\text{Reject } H_0 \text{ if } F_{STAT} > 3.07;$$
$$\text{otherwise, do not reject } H_0.$$

From Equation (10.9) on page 392,

$$F_{STAT} = \frac{S_1^2}{S_2^2} = \frac{210.2}{36.5} = 5.7589$$

Because $F_{STAT} = 5.7589 > 3.07$, you reject H_0. Using a 0.05 level of significance, you conclude that there is evidence that there is more variability in the final exam scores of students taking the introductory accounting course who are not majoring in accounting than for students taking the course who are majoring in accounting.

Problems for Section 10.4

LEARNING THE BASICS

10.36 Determine the upper-tail critical values of F in each of the following two-tail tests.
a. $\alpha = 0.10, n_1 = 16, n_2 = 21$
b. $\alpha = 0.05, n_1 = 16, n_2 = 21$
c. $\alpha = 0.01, n_1 = 16, n_2 = 21$

10.37 Determine the upper-tail critical value of F in each of the following one-tail tests.
a. $\alpha = 0.05, n_1 = 16, n_2 = 21$
b. $\alpha = 0.01, n_1 = 16, n_2 = 21$

10.38 The following information is available for two samples selected from independent normally distributed populations:

$$\text{Population A}: \quad n_1 = 25 \quad S_1^2 = 16$$
$$\text{Population B}: \quad n_2 = 25 \quad S_2^2 = 25$$

a. Which sample variance do you place in the numerator of F_{STAT}?
b. What is the value of F_{STAT}?

10.39 The following information is available for two samples selected from independent normally distributed populations:

$$\text{Population A}: \quad n_1 = 25 \quad S_1^2 = 161.9$$
$$\text{Population B}: \quad n_2 = 25 \quad S_2^2 = 133.7$$

What is the value of F_{STAT} if you are testing the null hypothesis $H_0: \sigma_1^2 = \sigma_2^2$?

10.40 In Problem 10.39, how many degrees of freedom are there in the numerator and denominator of the F test?

10.41 In Problems 10.39 and 10.40, what is the upper-tail critical value for F if the level of significance, α, is 0.05 and the alternative hypothesis is $H_1: \sigma_1^2 \neq \sigma_2^2$?

10.42 In Problems 10.39 through 10.41, what is your statistical decision?

10.43 The following information is available for two samples selected from independent but very right-skewed populations:

$$\text{Population A}: \quad n_1 = 16 \quad S_1^2 = 47.3$$
$$\text{Population B}: \quad n_2 = 13 \quad S_2^2 = 36.4$$

Should you use the F test to test the null hypothesis of equality of variances? Discuss.

10.44 In Problem 10.43, assume that two samples are selected from independent normally distributed populations.
a. At the 0.05 level of significance, is there evidence of a difference between σ_1^2 and σ_2^2?
b. Suppose that you want to perform a one-tail test. At the 0.05 level of significance, what is the upper-tail critical value of F to determine whether there is evidence that $\sigma_1^2 > \sigma_2^2$? What is your statistical decision?

APPLYING THE CONCEPTS

10.45 Shipments of meat, meat by-products, and other ingredients are mixed together in several filling lines at a pet food canning factory. Operations managers suspect that, although the mean amount filled per can of pet food is usually stable, the variability of the cans filled in line A is greater than that of line B. The following data from a sample of 8-ounce cans is as follows:

	Line A	Line B
$\bar{X}$	8.005	7.997
S	0.012	0.005
n	11	16

a. At the 0.05 level of significance, is there evidence that the variance in line A is greater than the variance in line B?
b. Interpret the p-value.
c. What assumption do you need to make in (a) about the two populations in order to justify your use of the F test?

SELF Test **10.46** The Computer Anxiety Rating Scale (CARS) measures an individual's level of computer anxiety, on a scale from 20 (no anxiety) to 100 (highest level of anxiety). Researchers at Miami University administered CARS to 172 business students. One of the objectives of the study was to determine whether there is a difference between the level of computer anxiety experienced by female students and male students. They found the following:

	Males	Females
$\bar{X}$	40.26	36.85
S	13.35	9.42
n	100	72

Source: Data extracted from T. Broome and D. Havelka, "Determinants of Computer Anxiety in Business Students," *The Review of Business Information Systems,* Spring 2002, 6(2), pp. 9–16.

a. At the 0.05 level of significance, is there evidence of a difference in the variability of the computer anxiety experienced by males and females?
b. Interpret the p-value.
c. What assumption do you need to make about the two populations in order to justify the use of the F test?
d. Based on (a) and (b), which t test defined in Section 10.1 should you use to test whether there is a significant difference in mean computer anxiety for female and male students?

10.47 A bank with a branch located in a commercial district of a city has the business objective of improving the process for serving customers during the noon-to-1 P.M. lunch period. To do so, the waiting time (defined as the time elapsed from when the customer enters the line until he or she reaches the teller window) needs to be shortened to increase customer satisfaction. A random sample of 15 customers is selected (and stored in **Bank1**), and the results (in minutes) are as follows:

4.21	5.55	3.02	5.13	4.77	2.34	3.54	3.20
4.50	6.10	0.38	5.12	6.46	6.19	3.79	

Suppose that another branch, located in a residential area, is also concerned with the noon-to-1 P.M. lunch period. A random sample of 15 customers is selected (and stored in **Bank2**), and the results (in minutes) are as follows:

9.66	5.90	8.02	5.79	8.73	3.82	8.01	8.35
10.49	6.68	5.64	4.08	6.17	9.91	5.47	

a. Is there evidence of a difference in the variability of the waiting time between the two branches? (Use $\alpha = 0.05$.)
b. Determine the p-value in (a) and interpret its meaning.
c. What assumption about the population distribution of the two banks is necessary in (a)? Is the assumption valid for these data?
d. Based on the results of (a), is it appropriate to use the pooled-variance t test to compare the means of the two branches?

10.48 An important feature of digital cameras is battery life, the number of shots that can be taken before the battery needs to be recharged. The file **DigitalCameras** contains battery life information for 29 subcompact cameras and 16 compact cameras. (Data extracted from "Digital Cameras," *Consumer Reports*, July 2009, pp. 28–29.)
a. Is there evidence of a difference in the variability of the battery life between the two types of digital cameras? (Use $\alpha = 0.05$.)
b. Determine the p-value in (a) and interpret its meaning.
c. What assumption about the population distribution of the two types of cameras is necessary in (a)? Is the assumption valid for these data?
d. Based on the results of (a), which t test defined in Section 10.1 should you use to compare the mean battery life of the two types of cameras?

10.49 Do young children use cell phones? Apparently so, according to a recent study (A. Ross, "Message to Santa; Kids Want a Phone," *Palm Beach Post*, December 16, 2008, pp. 1A, 4A), which stated that cell phone users under 12 years of age averaged 137 calls per month as compared to 231 calls per month for cell phone users 13 to 17 years of age. No sample sizes were reported. Suppose that the results were based on samples of 50 cell phone users in each group and that the sample standard deviation for cell phone users under 12 years of age was 51.7 calls per month and the sample standard deviation for cell phone users 13 to 17 years of age was 67.6 calls per month.
a. Using a 0.05 level of significance, is there evidence of a difference in the variances of cell phone usage between cell phone users under 12 years of age and cell phone users 13 to 17 years of age?
b. On the basis of the results in (a), which t test defined in Section 10.1 should you use to compare the means of the two groups of cell phone users? Discuss.

10.50 Is there a difference in the variation of the yield of five-year CDs at different times? The file **CD-FiveYear** contains the yields for a five-year certificate of deposit (CD) for 25 banks in the United States, as of March 29, 2010 and August 23, 2010. (Data extracted from **www.Bankrate.com**, March 29, 2010 and August 23, 2010.) At the 0.05 level of significance, is there evidence of a difference in the variance of the yield of five-year CDs on March 29, 2010 and August 23, 2010? Assume that the population yields are normally distributed.

USING STATISTICS @ BLK Beverages Revisited

In the Using Statistics scenario, you were the regional sales manager for BLK Beverages. You compared the sales volume of BLK Cola when the product is placed in the normal shelf location to the sales volume when the product is featured in a special end-aisle display. An experiment was performed in which 10 stores used the normal shelf location and 10 stores used the end-aisle displays. Using a t test for the difference between two means, you were able to conclude that the mean sales using end-aisle location are higher than the mean sales for the normal shelf location. A confidence interval allowed you to infer with 95% confidence that the end-aisle location sells, on average, 6.73 to 36.67 cases more than the normal shelf location. You also performed the F test for the difference between two variances to see if the store-to-store variability in sales in stores using the end-aisle location differed from the store-to-store variability in sales in stores using the normal shelf location. You concluded that there was no significant difference in the variability of the sales of cola for the two display locations. As regional sales manager, your next step in increasing sales is to convince more stores to use the special end-aisle display.

SUMMARY

In this chapter, you were introduced to a variety of two-sample tests. For situations in which the samples are independent, you learned statistical test procedures for analyzing possible differences between means, variances, and proportions. In addition, you learned a test procedure that is frequently used when analyzing differences between the means of two related samples. Remember that you need to select the test that is most appropriate for a given set of conditions and to critically investigate the validity of the assumptions underlying each of the hypothesis-testing procedures.

Table 10.7 provides a list of topics covered in this chapter. The roadmap in Figure 10.15 illustrates the steps needed in determining which two-sample test of hypothesis to use. The following are the questions you need to consider:

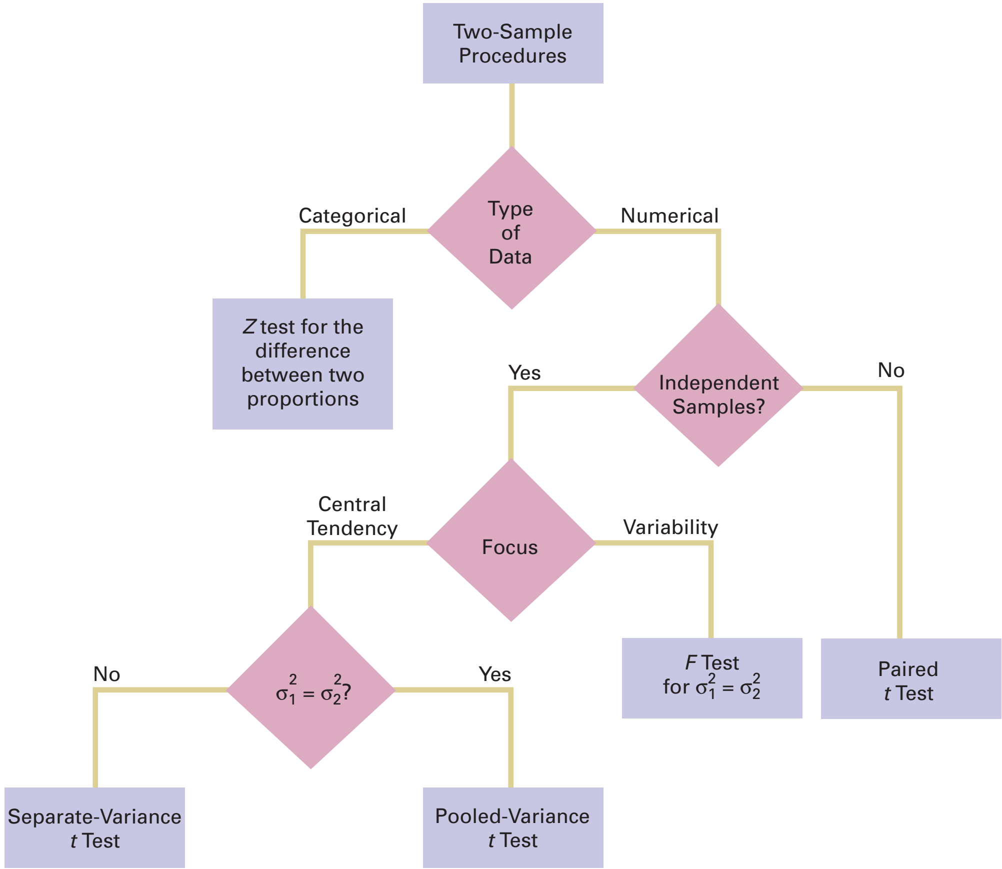
1. What type of data do you have? If you are dealing with categorical variables, use the Z test for the difference between two proportions. (This test assumes independent samples.)
2. If you have a numerical variable, determine whether you have independent samples or related samples. If you have related samples, and you can assume approximate normality, use the paired t test.
3. If you have independent samples, is your focus on variability or central tendency? If the focus is on variability, and you can assume approximate normality, use the F test.
4. If your focus is central tendency and you can assume approximate normality, determine whether you can assume that the variances of the two populations are equal. (This assumption can be tested using the F test.)
5. If you can assume that the two populations have equal variances, use the pooled-variance t test. If you cannot assume that the two populations have equal variances, use the separate-variance t test.

TABLE 10.7

Summary of Topics in Chapter 10

	Types of Data	
Type of Analysis	**Numerical**	**Categorical**
Comparing two populations	t tests for the difference in the means of two independent populations (Section 10.1)	Z test for the difference between two proportions (Section 10.3)
	Paired t test (Section 10.2)	
	F test for the ratio of two variances (Section 10.4)	

FIGURE 10.15
Roadmap for selecting a two-sample test of hypothesis



KEY EQUATIONS

Pooled-Variance *t* Test for the Difference Between Two Means

$$t_{STAT} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \textbf{(10.1)}$$

Confidence Interval Estimate for the Difference in the Means of Two Independent Populations

$$(\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \quad \textbf{(10.2)}$$

or

$$(\bar{X}_1 - \bar{X}_2) - t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \le \mu_1 - \mu_2 \le (\bar{X}_1 - \bar{X}_2) + t_{\alpha/2}\sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Separate-Variance *t* Test for the Difference Between Two Means

$$t_{STAT} = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \quad \textbf{(10.3)}$$

Computing Degrees of Freedom in the Separate-Variance *t* Test

$$V = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{\left(\frac{S_1^2}{n_1}\right)^2}{n_1 - 1} + \frac{\left(\frac{S_2^2}{n_2}\right)^2}{n_2 - 1}} \quad \textbf{(10.4)}$$

Paired *t* Test for the Mean Difference

$$t_{STAT} = \frac{\bar{D} - \mu_D}{\frac{S_D}{\sqrt{n}}} \quad \textbf{(10.5)}$$

Confidence Interval Estimate for the Mean Difference

$$\overline{D} \pm t_{\alpha/2}\frac{S_D}{\sqrt{n}} \tag{10.6}$$

or

$$\overline{D} - t_{\alpha/2}\frac{S_D}{\sqrt{n}} \le \mu_D \le \overline{D} + t_{\alpha/2}\frac{S_D}{\sqrt{n}}$$

***Z* Test for the Difference Between Two Proportions**

$$Z_{STAT} = \frac{(p_1 - p_2) - (\pi_1 - \pi_2)}{\sqrt{\overline{p}(1 - \overline{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \tag{10.7}$$

Confidence Interval Estimate for the Difference Between Two Proportions

$$(p_1 - p_2) \pm Z_{\alpha/2}\sqrt{\left(\frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1 - p_2)}{n_2}\right)} \tag{10.8}$$

or

$$(p_1 - p_2) - Z_{\alpha/2}\sqrt{\frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1 - p_2)}{n_1}} \le (\pi_1 - \pi_2)$$
$$\le (p_1 - p_2) + Z_{\alpha/2}\sqrt{\frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1 - p_2)}{n_2}}$$

***F* Test Statistic for Testing the Ratio of Two Variances**

$$F_{STAT} = \frac{S_1^2}{S_2^2} \tag{10.9}$$

KEY TERMS

F distribution 392
F test for the ratio of two variances 392
matched samples 377
paired *t* test for the mean difference 378
pooled-variance *t* test 366
repeated measurements 377
robust 369
separate-variance *t* test 372
two-sample test 366
Z test for the difference between two proportions 386

CHAPTER REVIEW PROBLEMS

CHECKING YOUR UNDERSTANDING

10.51 What are some of the criteria used in the selection of a particular hypothesis-testing procedure?

10.52 Under what conditions should you use the pooled-variance *t* test to examine possible differences in the means of two independent populations?

10.53 Under what conditions should you use the *F* test to examine possible differences in the variances of two independent populations?

10.54 What is the distinction between two independent populations and two related populations?

10.55 What is the distinction between repeated measurements and matched items?

10.56 When you have two independent populations, explain the similarities and differences between the test of hypothesis for the difference between the means and the confidence interval estimate for the difference between the means.

10.57 Under what conditions should you use the paired *t* test for the mean difference between two related populations?

APPLYING THE CONCEPTS

10.58 The per-store daily customer count (i.e., the mean number of customers in a store in one day) for a nationwide convenience store chain that operates nearly 10,000 stores has been steady, at 900, for some time. To increase the customer count, the chain is considering cutting prices for coffee beverages. The small size will now be \$0.59 instead of \$0.99, and medium size will be \$0.69 instead of \$1.19. Even with this reduction in price, the chain will have a 40% gross margin on coffee. The question to be determined is how much to cut prices to increase the daily customer count without reducing the gross margin on coffee sales too much. The chain decides to carry out an experiment in a sample of 30 stores where customer counts have been running almost exactly at the national average of 900. In 15 of the stores, the price of a small coffee will now be \$0.59 instead of \$0.99, and in 15 other stores, the price of a small coffee will now be \$0.79. After four weeks, the 15 stores that priced the small coffee at \$0.59 had a mean daily customer count of 964 and a standard deviation of 88, and the 15 stores that priced the small coffee at \$0.79 had a mean daily customer count of 941 and a standard deviation of 76. Analyze these data (use the 0.05 level of significance) and answer the following questions.

a. Does reducing the price of a small coffee to either \$0.59 or \$0.79 increase the mean per-store daily customer count?

b. If reducing the price of a small coffee to either \$0.59 or \$0.79 increases the mean per-store daily customer count, is there any difference in the mean per-store daily customer count between stores in which a small coffee was priced at \$0.59 and stores in which a small coffee was priced at \$0.79?

c. What price do you recommend that a small coffee should be sold for?

10.59 A study conducted in March 2009 found that about half of U.S. adults trusted the U.S. government more than U.S. business to solve the economic problems of the United States. However, when the population is subdivided by political party affiliation, the results are very different. The study showed that 72% of Democrats trusted the government more, but only 29% of Republicans trusted the government more. Suppose that you are in charge of updating the study. You will take a national sample of Democrats and a national sample of Republicans and then try to use the results to show statistical evidence that the proportion of Democrats trusting the government more than business is greater than the proportion of Republicans trusting the government more than business.

a. What are the null and alternative hypotheses?

b. What is a Type I error in the context of this study?

c. What is a Type II error in the context of this study?

10.60 The American Society for Quality (ASQ) conducted a salary survey of all its members. ASQ members work in all areas of manufacturing and service-related institutions, with a common theme of an interest in quality. Two job titles are black belt and green belt. (See Section 17.8, for a description of these titles in a Six Sigma quality improvement initiative.) Descriptive statistics concerning salaries for these two job titles are given in the following table:

Job Title	Sample Size	Mean	Standard Deviation
Black Belt	219	87,342	20,955
Green belt	34	65,679	18,137

Source: Data extracted from J. D. Conklin, "Salary Survey: Holding Steady," *Quality Progress,* December 2009, pp. 20–53.

a. Using a 0.05 level of significance, is there a difference in the variability of salaries between black belts and green belts?

b. Based on the result of (a), which *t* test defined in Section 10.1 is appropriate for comparing mean salaries?

c. Using a 0.05 level of significance, is there a difference between the mean salary of black belts and green belts?

10.61 Do male and female students study the same amount per week? In 2007, 58 sophomore business students were surveyed at a large university that has more than 1,000 sophomore business students each year. The file **StudyTime** contains the gender and the number of hours spent studying in a typical week for the sampled students.

a. At the 0.05 level of significance, is there a difference in the variance of the study time for male students and female students?

b. Using the results of (a), which *t* test is appropriate for comparing the mean study time for male and female students?

c. At the 0.05 level of significance, conduct the test selected in (b).

d. Write a short summary of your findings.

10.62 Two professors wanted to study how students from their two universities compared in their capabilities of using Excel spreadsheets in undergraduate information systems courses. (Data extracted from H. Howe and M. G. Simkin, "Factors Affecting the Ability to Detect Spreadsheet Errors," *Decision Sciences Journal of Innovative Education*, January 2006, pp. 101–122.) A comparison of the student demographics was also performed. One school is a state university in the western United States, and the other school is a state university in the eastern United States. The following table contains information regarding the ages of the students:

School	Sample Size	Mean	Standard Deviation
Western	93	23.28	6.29
Eastern	135	21.16	1.32

a. Using a 0.01 level of significance, is there evidence of a difference in the variances of the age of students at the western school and at the eastern school?

b. Discuss the practical implications of the test performed in (a). Address, specifically, the impact equal (or unequal) variances in age has on teaching an undergraduate information systems course.

c. To test for a difference in the mean age of students, is it most appropriate to use the pooled-variance *t* test or the separate-variance *t* test?

School	Sample Size	Mean	Standard Deviation
Western	93	2.6	2.4
Eastern	135	4.0	2.1

The following table contains information regarding the years of spreadsheet usage of the students:

d. Using a 0.01 level of significance, is there evidence of a difference in the variances of the years of spreadsheet usage of students at the western school and at the eastern school?

e. Based on the results of (d), use the most appropriate test to determine, at the 0.01 level of significance, whether there is evidence of a difference in the mean years of spreadsheet usage of students at the western school and at the eastern school.

10.63 The file **Restaurants** contains the ratings for food, décor, service, and the price per person for a sample of 50 restaurants located in a city and 50 restaurants located in a suburb. Completely analyze the differences between city and suburban restaurants for the variables food rating, décor rating, service rating, and cost per person, using $\alpha = 0.05$.

Source: Data extracted from *Zagat Survey 2010: New York City Restaurants* and *Zagat Survey 2009–2010: Long Island Restaurants.*

10.64 A computer information systems professor is interested in studying the amount of time it takes students enrolled in the introduction to computers course to write and run a program in Visual Basic. The professor hires you to analyze the following results (in minutes) from a random sample of nine students (the data are stored in the **VB** file):

10 13 9 15 12 13 11 13 12

a. At the 0.05 level of significance, is there evidence that the population mean amount is greater than 10 minutes? What will you tell the professor?

b. Suppose that the professor, when checking her results, realizes that the fourth student needed 51 minutes rather than the recorded 15 minutes to write and run the Visual Basic program. At the 0.05 level of significance, reanalyze the question posed in (a), using the revised data. What will you tell the professor now?

c. The professor is perplexed by these paradoxical results and requests an explanation from you regarding the justification for the difference in your findings in (a) and (b). Discuss.

d. A few days later, the professor calls to tell you that the dilemma is completely resolved. The original number 15 (the fourth data value) was correct, and therefore your findings in (a) are being used in the article she is writing for a computer journal. Now she wants to hire you to compare the results from that group of introduction to computers students against those from a sample of 11 computer majors in order to determine whether there is evidence that computer majors can write a Visual Basic program in less time than introductory students. For the computer majors, the sample mean is 8.5 minutes, and the sample standard deviation is 2.0 minutes. At the 0.05 level of significance, completely analyze these data. What will you tell the professor?

e. A few days later, the professor calls again to tell you that a reviewer of her article wants her to include the p-value for the "correct" result in (a). In addition, the professor inquires about an unequal-variances problem, which the reviewer wants her to discuss in her article. In your own words, discuss the concept of p-value and also describe the unequal-variances problem. Then, determine the p-value in (a) and discuss whether the unequal-variances problem had any meaning in the professor's study.

10.65 An article (A. Jennings, "What's Good for a Business Can Be Hard on Friends," *The New York Times*, August 4, 2007, pp. C1–C2) reported that according to a poll, the mean number of cell phone calls per month was 290 for 18- to 24-year-olds and 194 for 45- to 54-year-olds, whereas the mean number of text messages per month was 290 for 18- to 24-year-olds and 57 for 45- to 54-year-olds. Suppose that the poll was based on a sample of 100 18- to 24-year-olds and 100 45- to 54-year-olds and that the standard deviation of the number of cell phone calls per month was 100 for 18- to 24-year-olds and 90 for 45- to 54-year-olds, whereas the standard deviation of the number of text messages per month was 90 for 18- to 24-year-olds and 77 for 45- to 54-year-olds. Assume a level of significance of 0.05.

a. Is there evidence of a difference in the variances of the number of cell phone calls per month for 18- to 24-year-olds and for 45- to 54-year-olds?

b. Is there evidence of a difference in the mean number of cell phone calls per month for 18- to 24-year-olds and for 45- to 54-year-olds?

c. Construct and interpret a 95% confidence interval estimate for the difference in the mean number of cell phone calls per month for 18- to 24-year-olds and 45- to 54-year-olds.

d. Is there evidence of a difference in the variances of the number of text messages per month for 18- to 24-year-olds and 45- to 54-year-olds?

e. Is there evidence of a difference in the mean number of text messages per month for 18- to 24-year-olds and 45- to 54-year-olds?

f. Construct and interpret a 95% confidence interval estimate for the difference in the mean number of text messages per month for 18- to 24-year-olds and 45- to 54-year-olds.

g. Based on the results of (a) through (f), what conclusions can you make concerning cell phone and text message usage between 18- to 24-year-olds and 45- to 54-year-olds?

10.66 The lengths of life (in hours) of a sample of 40 100-watt light bulbs produced by manufacturer A and a sample of 40 100-watt light bulbs produced by manufacturer B are stored in **Bulbs**. Completely analyze the differences between the lengths of life of the bulbs produced by the two manufacturers. (Use $\alpha = 0.05$.)

10.67 A hotel manager looks to enhance the initial impressions that hotel guests have when they check in. Contributing to initial impressions is the time it takes to deliver a guest's luggage to the room after check-in. A random sample of 20 deliveries on a particular day were selected in Wing A of the hotel, and a random sample of 20 deliveries were selected in Wing B. The results are stored in **Luggage**. Analyze the data and determine whether there is a difference in the mean delivery time in the two wings of the hotel. (Use $\alpha = 0.05$.)

10.68 According to Census estimates, there are about 20 million children between 8 and 12 years old (referred to as *tweens*) in the United States in 2009. A recent survey of 1,223 8- to 12-year-old children (S. Jayson, "It's Cooler Than Ever to Be a Tween," *USA Today*, February 4, 2009, pp. 1A, 2A) reported the following results. Suppose the survey was based on 600 boys and 623 girls.

What Tweens Did in the Past Week	Boys	Girls
Played a game on a video game system	498	243
Read a book for fun	276	324
Gave product advice to parents	186	181
Shopped at a mall	144	262

For *each type of activity*, determine whether there is a difference between boys and girls at the 0.05 level of significance.

10.69 The manufacturer of Boston and Vermont asphalt shingles knows that product weight is a major factor in the customer's perception of quality. Moreover, the weight represents the amount of raw materials being used and is therefore very important to the company from a cost standpoint. The last stage of the assembly line packages the shingles before they are placed on wooden pallets. Once a pallet is full (a pallet for most brands holds 16 squares of shingles), it is weighed, and the measurement is recorded. The file **Pallet** contains the weight (in pounds) from a sample of 368 pallets of Boston shingles and 330 pallets of Vermont shingles. Completely analyze the differences in the weights of the Boston and Vermont shingles, using $\alpha = 0.05$.

10.70 The manufacturer of Boston and Vermont asphalt shingles provides its customers with a 20-year warranty on most of its products. To determine whether a shingle will last as long as the warranty period, the manufacturer conducts accelerated-life testing. Accelerated-life testing exposes the shingle to the stresses it would be subject to in a lifetime of normal use in a laboratory setting via an experiment that takes only a few minutes to conduct. In this test, a shingle is repeatedly scraped with a brush for a short period of time, and the shingle granules removed by the brushing are weighed (in grams). Shingles that experience low amounts of granule loss are expected to last longer in normal use than shingles that experience high amounts of granule loss. In this situation, a shingle should experience no more than 0.8 grams of granule loss if it is expected to last the length of the warranty period. The file **Granule** contains a sample of 170 measurements made on the company's Boston shingles and 140 measurements made on Vermont shingles. Completely analyze the differences in the granule loss of the Boston and Vermont shingles, using $\alpha = 0.05$.

10.71 There are a tremendous number of mutual funds from which an investor can choose. Each mutual fund has its own mix of different types of investments. The data in **BestFunds** present the 3-year annualized return, 5-year annualized return, 10-year annualized return, and expense ratio (in %) for the 10 best mutual funds according to the *U.S. News & World Report* score for large cap value and large cap growth mutual funds. (Data extracted from K. Shinkle, "The Best Funds for the Long Term, *U.S. News & World Report*, Summer 2010, pp. 52–56.) Analyze the data and determine whether any differences exist between large cap value and large cap growth mutual funds. (Use the 0.05 level of significance.)

REPORT WRITING EXERCISE

10.72 Referring to the results of Problems 10.69 and 10.70 concerning the weight and granule loss of Boston and Vermont shingles, write a report that summarizes your conclusions.

TEAM PROJECT

The file **Bond Funds** contains information regarding eight variables from a sample of 184 bond mutual funds:

Type—Type of bonds comprising the bond fund (intermediate government or short-term corporate)
Assets—In millions of dollars
Fees—Sales charges (no or yes)
Expense ratio—Ratio of expenses to net assets, in percentage
Return 2009—Twelve-month return in 2009
Three-year return—Annualized return, 2007–2009
Five-year return—Annualized return, 2005–2009
Risk—Risk-of-loss factor of the mutual fund (below average, average, or above average)

10.73 Completely analyze the differences between bond mutual funds without fees and bond mutual funds with fees in terms of 2009 return, three-year return, five-year return, and expense ratio. Write a report summarizing your findings.

10.74 Completely analyze the difference between intermediate government bond mutual funds and short-term corporate bond mutual funds in terms of 2009 return, three-year return, five-year return, and expense ratio. Write a report summarizing your findings.

STUDENT SURVEY DATABASE

10.75 Problem 1.27 on page 14 describes a survey of 62 undergraduate students (stored in **UndergradSurvey**).

a. At the 0.05 level of significance, is there evidence of a difference between males and females in grade point average, expected starting salary, number of social networking sites registered for, age, spending on textbooks and supplies, text messages sent in a week, and the wealth needed to feel rich?

b. At the 0.05 level of significance, is there evidence of a difference between students who plan to go to graduate school and those who do not plan to go to graduate school in grade point average, expected starting salary, number of social networking sites registered for, age, spending on textbooks and supplies, text messages sent in a week, and the wealth needed to feel rich?

10.76 Problem 1.27 on page 14 describes a survey of 62 undergraduate students (stored in **UndergradSurvey**).

a. Select a sample of undergraduate students at your school and conduct a similar survey for them.

b. For the data collected in (a), repeat (a) and (b) of Problem 10.75.

c. Compare the results of (b) to those of Problem 10.75.

10.77 Problem 1.28 on page 15 describes a survey of 44 MBA students (stored in **GradSurvey**). For these data, at the 0.05 level of significance, is there evidence of a difference between males and females in age, undergraduate grade point average, graduate grade point average, expected salary upon graduation, spending on textbooks and supplies, text messages sent in a week, and the wealth needed to feel rich?

10.78 Problem 1.28 on page 15 describes a survey of 44 MBA students (stored in **GradSurvey**).

a. Select a sample of graduate students in your MBA program and conduct a similar survey for those students.

b. For the data collected in (a), repeat Problem 10.77.

c. Compare the results of (b) to those of Problem 10.77.

MANAGING ASHLAND MULTICOMM SERVICES

AMS communicates with customers who subscribe to cable television services through a special secured email system that sends messages about service changes, new features, and billing information to in-home digital set-top boxes for later display. To enhance customer service, the operations department established the business objective of reducing the amount of time to fully update each subscriber's set of messages. The department selected two candidate messaging systems and conducted an experiment in which 30 randomly chosen cable subscribers were assigned one of the two systems (15 assigned to each system). Update times were measured, and the results are organized in Table AMS10.1 (and stored in **AMS10**).

EXERCISES

1. Analyze the data in Table AMS10.1 and write a report to the computer operations department that indicates your findings. Include an appendix in which you discuss the reason you selected a particular statistical test to compare the two independent groups of callers.

2. Suppose that instead of the research design described in the case, there were only 15 subscribers sampled, and the update process for each subscriber e-mail was measured for each of the two messaging systems. Suppose the results were organized in Table AMS10.1—making each row in the table a pair of values for an individual subscriber. Using these suppositions, reanalyze the Table AMS10.1 data and write a report for presentation to the team that indicates your findings.

TABLE AMS10.1

Download Time for Two Different E-mail Interfaces

Email Interface 1	Email Interface 2
4.13	3.71
3.75	3.89
3.93	4.22
3.74	4.57
3.36	4.24
3.85	3.90
3.26	4.09
3.73	4.05
4.06	4.07
3.33	3.80
3.96	4.36
3.57	4.38
3.13	3.49
3.68	3.57
3.63	4.74

DIGITAL CASE

Apply your knowledge about hypothesis testing in this Digital Case, which continues the cereal-fill packaging dispute Digital Case from Chapters 7 and 9.

Even after the recent public experiment about cereal box weights, Consumers Concerned About Cereal Cheaters (CCACC) remains convinced that Oxford Cereals has misled the public. The group has created and circulated **MoreCheating.pdf,** a document in which it claims that cereal boxes produced at Plant Number 2 in Springville weigh less than the claimed mean of 368 grams. Review this document and then answer the following questions:

1. Do the CCACC's results prove that there is a statistically significant difference in the mean weights of cereal boxes produced at Plant Numbers 1 and 2?

2. Perform the appropriate analysis to test the CCACC's hypothesis. What conclusions can you reach based on the data?

REFERENCES

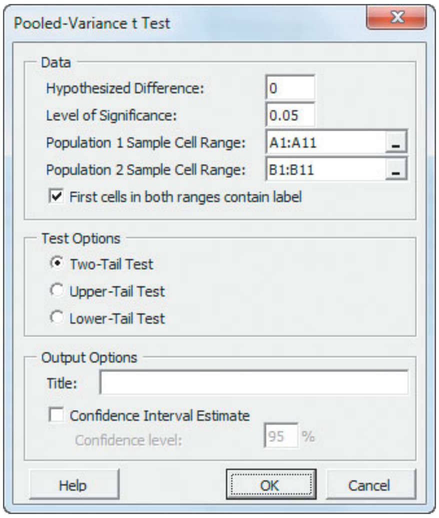
1. Conover, W. J., *Practical Nonparametric Statistics*, 3rd ed. (New York: Wiley, 2000).
2. Daniel, W., *Applied Nonparametric Statistics*, 2nd ed. (Boston: Houghton Mifflin, 1990).
3. *Microsoft Excel 2010* (Redmond, WA: Microsoft Corp., 2010).
4. *Minitab Release 16* (State College, PA: Minitab, Inc., 2010).
5. Satterthwaite, F. E., "An Approximate Distribution of Estimates of Variance Components," *Biometrics Bulletin*, 2(1946): 110–114.
6. Snedecor, G. W., and W. G. Cochran, *Statistical Methods*, 8th ed. (Ames, IA: Iowa State University Press, 1989).
7. Winer, B. J., D. R. Brown, and K. M. Michels, *Statistical Principles in Experimental Design*, 3rd ed. (New York: McGraw-Hill, 1989).

EG10.1 COMPARING the MEANS of TWO INDEPENDENT POPULATIONS

Pooled-Variance *t* Test for the Difference Between Two Means

PHStat2 Use **Pooled-Variance t Test** to perform the pooled-variance *t* test. For example, to perform the Figure 10.3 pooled-variance *t* test for the BLK Cola data shown on page 365, open to the **DATA worksheet** of the **COLA workbook**. Select **PHStat ➔ Two-Sample Tests (Unsummarized Data) ➔ Pooled-Variance t Test**. In the procedure's dialog box (shown below):

1. Enter **0** as the **Hypothesized Difference**.
2. Enter **0.05** as the **Level of Significance**.
3. Enter **A1:A11** as the **Population 1 Sample Cell Range**.
4. Enter **B1:B11** as the **Population 2 Sample Cell Range**.
5. Check **First cells in both ranges contain label**.
6. Click **Two-Tail Test**.
7. Enter a **Title** and click **OK**.



For problems that use summarized data, select **PHStat ➔ Two-Sample Tests (Summarized Data) ➔ Pooled-Variance t Test**. In that procedure's dialog box, enter the hypothesized difference and level of significance, as well as the sample size, sample mean, and sample standard deviation for each sample.

In-Depth Excel Use the **COMPUTE worksheet** of the **Pooled-Variance T workbook**, shown in Figure 10.3 on page 368, as a template for performing the two-tail pooled-variance *t* test. The worksheet contains data and formulas to use the unsummarized data for the BLK Cola example. In cell B24 and B25, respectively, the worksheet uses the expressions **-TINV(*level of significance*, *degrees of freedom*)** and **TINV(*level of significance*, *degrees of freedom*)** to compute the lower and upper critical values. In cell B26, **TDIST(*absolute value of the t test statistic*, *degrees of freedom*, 2)** computes the *p*-value.

For other problems, use the COMPUTE worksheet with either unsummarized or summarized data. For unsummarized data, keep the formulas that calculate the sample size, sample mean, and sample standard deviation in cell ranges B7:B9 and B11:B13 and change the data in columns A and B in the **DATACOPY worksheet**. For summarized data, replace the formulas in cell ranges B7:B9 and B11:B13 with the sample statistics and ignore the DATACOPY worksheet.

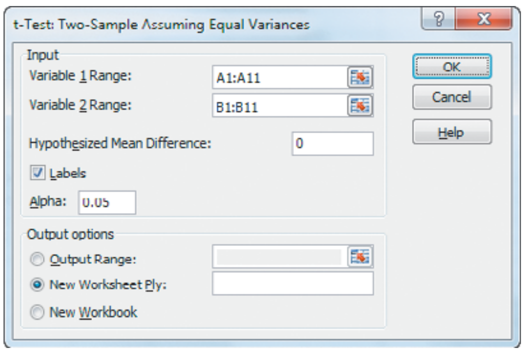
Use the similar **COMPUTE_LOWER** or **COMPUTE_UPPER worksheets** in the same workbook as templates for performing one-tail pooled-variance *t* tests. These worksheets can also use either unsummarized or summarized data.

Analysis ToolPak Use **t-Test: Two-Sample Assuming Equal Variances** to perform the pooled-variance *t* test for unsummarized data. For example, to create results equivalent to those in the Figure 10.3 pooled-variance *t* test for the BLK Cola example on page 368, open to the **DATA worksheet** of the **COLA workbook** and:

1. Select **Data ➔ Data Analysis**.
2. In the Data Analysis dialog box, select **t-Test: Two-Sample Assuming Equal Variances** from the **Analysis Tools** list and then click **OK**.

In the procedure's dialog box (shown below):

3. Enter **A1:A11** as the **Variable 1 Range** and enter **B1:B11** as the **Variable 2 Range**.
4. Enter **0** as the **Hypothesized Mean Difference**.
5. Check **Labels** and enter **0.05** as **Alpha**.
6. Click **New Worksheet Ply**.
7. Click **OK**.



Results (shown below) appear in a new worksheet that contains both two-tail and one-tail test critical values and *p*-values. Unlike Figure 10.3, only the positive (upper) critical value is listed for the two-tail test.

	A	B	C
1	t-Test: Two-Sample Assuming Equal Variances		
2			
3		*Normal*	*EndAisle*
4	Mean	50.3	72
5	Variance	350.6778	157.3333
6	Observations	10	10
7	Pooled Variance	254.0056	
8	Hypothesized Mean Difference	0	
9	df	18	
10	t Stat	-3.04455	
11	P(T<=t) one-tail	0.003487	
12	t Critical one-tail	1.734064	
13	P(T<=t) two-tail	0.006975	
14	t Critical two-tail	2.100922	

Confidence Interval Estimate for the Difference Between Two Means

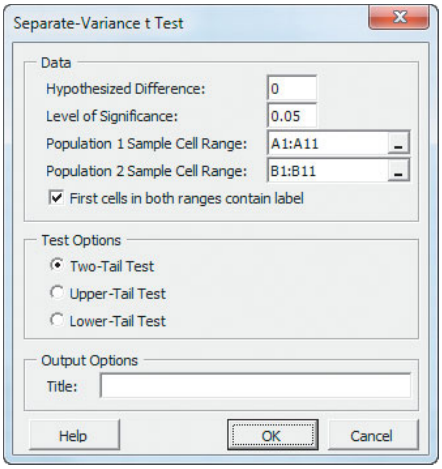
PHStat2 Use the *PHStat2* instructions for the pooled-variance *t* test. In step 7, also check **Confidence Interval Estimate** and enter a **Confidence Level** in its box, in addition to entering a **Title** and clicking **OK**.

In-Depth Excel Use the *In-Depth Excel* instructions for the pooled-variance *t* test. The worksheets in the **Pooled-Variance T workbook** include a confidence interval estimate for the difference between two means in the cell range D3:E16.

t Test for the Difference Between Two Means Assuming Unequal Variances

PHStat2 Use **Separate-Variance t Test** to perform this *t* test. For example, to perform the Figure 10.7 separate-variance *t* test for the BLK Cola data on page 374, open to the **DATA worksheet** of the **COLA workbook**. Select **PHStat ➔ Two-Sample Tests (Unsummarized Data) ➔ Separate-Variance t Test**. In the procedure's dialog box (shown at the top of the right column):

1. Enter **0** as the **Hypothesized Difference**.
2. Enter **0.05** as the **Level of Significance**.
3. Enter **A1:A11** as the **Population 1 Sample Cell Range**.
4. Enter **B1:B11** as the **Population 2 Sample Cell Range**.
5. Check **First cells in both ranges contain label**.
6. Click **Two-Tail Test**.
7. Enter a **Title** and click **OK**.



For problems that use summarized data, select **PHStat ➔ Two-Sample Tests (Summarized Data) ➔ Separate-Variance t Test**. In that procedure's dialog box, enter the hypothesized difference and the level of significance, as well as the sample size, sample mean, and sample standard deviation for each group.

In-Depth Excel Use the **COMPUTE worksheet** of the **Separate-Variance T workbook**, shown in Figure 10.7 on page 374, as a template for performing the two-tail separate-variance *t* test. The worksheet contains data and formulas to use the unsummarized data for the BLK Cola example. In cells B25 and B26, respectively, **-TINV(*level of significance*, *degrees of freedom*)** and **TINV(*level of significance*, *degrees of freedom*)** computes the lower and upper critical values. In cell B27, the worksheet uses **TDIST(*absolute value of the t test statistic*, *degrees of freedom*, 2)** to compute the *p*-value.

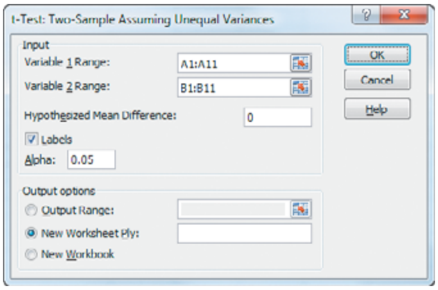
For other problems, use the COMPUTE worksheet with either unsummarized or summarized data. For unsummarized data, keep the formulas that calculate the sample size, sample mean, and sample standard deviation in cell ranges B7:B9 and B11:B13 and change the data in columns A and B in the **DATACOPY worksheet**. For summarized data, replace the formulas in cell ranges B7:B9 and B11:B13 with the sample statistics and ignore the DATACOPY worksheet. Use the similar **COMPUTE_LOWER** or **COMPUTE_UPPER worksheets** in the same workbook as templates for performing one-tail *t* tests.

Analysis ToolPak Use **t-Test: Two-Sample Assuming Unequal Variances** to perform the separate-variance *t* test for unsummarized data. For example, to create results equivalent to those in the Figure 10.7 separate-variance *t* test for the BLK Cola data on page 374, open to the **DATA worksheet** of the **COLA workbook** and:

1. Select **Data ➔ Data Analysis**.
2. In the Data Analysis dialog box, select **t-Test: Two-Sample Assuming Unequal Variances** from the **Analysis Tools** list and then click **OK**.

In the procedure's dialog box (shown below):

3. Enter **A1:A11** as the **Variable 1 Range** and enter **B1:B11** as the **Variable 2 Range**.
4. Enter **0** as the **Hypothesized Mean Difference**.
5. Check **Labels** and enter **0.05** as **Alpha**.
6. Click **New Worksheet Ply**.
7. Click **OK**.



Results (shown below) appear in a new worksheet that contains both two-tail and one-tail test critical values and *p*-values. Unlike Figure 10.7, only the positive (upper) critical value is listed for the two-tail test. Because the Analysis ToolPak uses table lookups to approximate the critical values and the *p*-value, the results will differ slightly from the values shown in Figure 10.7.

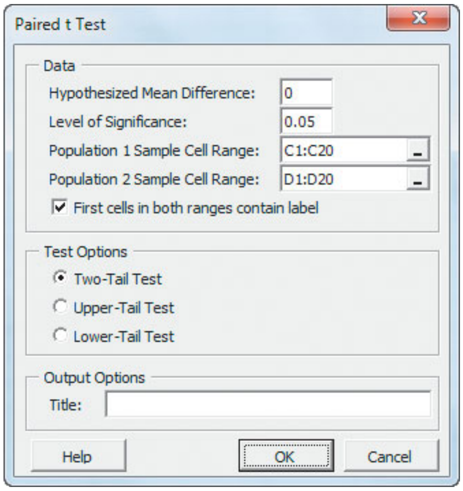
	A	B	C
1	t-Test: Two-Sample Assuming Unequal Variances		
2			
3		*Normal*	*EndAisle*
4	Mean	50.3	72
5	Variance	350.6778	157.3333
6	Observations	10	10
7	Hypothesized Mean Difference	0	
8	df	16	
9	t Stat	-3.04455	
10	P(T<=t) one-tail	0.003863	
11	t Critical one-tail	1.745884	
12	P(T<=t) two-tail	0.007726	
13	t Critical two-tail	2.119905	

EG10.2 COMPARING the MEANS of TWO RELATED POPULATIONS

Paired t Test

PHStat2 Use **Paired t Test** to perform the paired *t* test. For example, to perform the Figure 10.9 paired *t* test for the textbook price data on page 381, open to the **DATA worksheet** of the **BookPrices workbook**. Select **PHStat → Two-Sample Tests (Unsummarized Data) → Paired t Test**. In the procedure's dialog box (shown in the right column):

1. Enter **0** as the **Hypothesized Mean Difference**.
2. Enter **0.05** as the **Level of Significance**.
3. Enter C**1:C20** as the **Population 1 Sample Cell Range**.
4. Enter **D1:D20** as the **Population 2 Sample Cell Range**.
5. Check **First cells in both ranges contain label**.
6. Click **Two-Tail Test**.
7. Enter a **Title** and click **OK**.



The procedure creates two worksheets, one of which is similar to the PtCalcs worksheet discussed in the following *In-Depth Excel* section. For problems that use summarized data, select **PHStat → Two-Sample Tests (Summarized Data) → Paired t Test**. In that procedure's dialog box, enter the hypothesized mean difference and the level of significance, as well as the sample size, sample mean, and sample standard deviation for each sample.

In-Depth Excel Use the **COMPUTE** and **PtCalcs worksheets** of the **Paired T workbook**, as a template for performing the two-tail paired *t* test. The PtCalcs worksheet contains the differences and other intermediate calculations that allow the COMPUTE worksheet, shown in Figure 10.9 on page 381, to compute the sample size, $\overline{D}$, and S_D.

The COMPUTE and PtCalcs worksheets contain the data and formulas for the unsummarized data for the textbook prices example. In cells B16 and B17, respectively, the COMPUTE worksheet uses **-TINV(*level of significance*, *degrees of freedom*)** and **TINV(*level of significance*, *degrees of freedom*)** to compute the lower and upper critical values. In cell B18, the worksheet uses **TDIST(*absolute value of the t test statistic*, *degrees of freedom*, 2)** to compute the *p*-value.

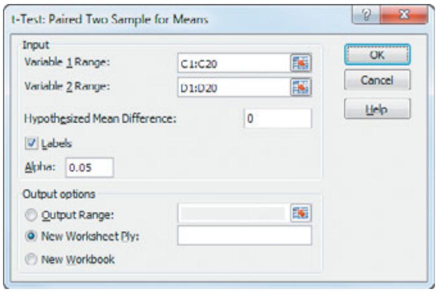
For other problems, paste the unsummarized data into columns A and B of the PtCalcs worksheet. For sample sizes greater than 19, select the cell range C20:D20 and copy the formulas in those cells down through the last data row. For sample sizes less than 19, delete the column C and D formulas for which there are no column A and B values. If you know the sample size, $\overline{D}$, and S_D values, you can ignore the PtCalcs worksheet and enter the values in cells B8, B9, and B11 of the COMPUTE worksheet, overwriting the formulas that those cells contain. Use the **COMPUTE_LOWER** or **COMPUTE_UPPER worksheets** in the same workbook as templates for performing one-tail tests.

Analysis ToolPak Use **t-Test: Paired Two Sample for Means** to perform the paired *t* test for unsummarized data. For example, to create results equivalent to those in the Figure 10.9 paired *t* test for the textbook price data on page 381, open to the **DATA worksheet** of the **BookPrices workbook** and:

1. Select **Data ➔ Data Analysis**.
2. In the Data Analysis dialog box, select **t-Test: Paired Two Sample for Means** from the **Analysis Tools** list and then click **OK**.

In the procedure's dialog box (shown below):

3. Enter **C1:C20** as the **Variable 1 Range** and enter **D1:D20** as the **Variable 2 Range**.
4. Enter **0** as the **Hypothesized Mean Difference**.
5. Check **Labels** and enter **0.05** as **Alpha**.
6. Click **New Worksheet Ply**.
7. Click **OK**.



Results (shown below) appear in a new worksheet that contains both two-tail and one-tail test critical values and *p*-values. Unlike Figure 10.9, only the positive (upper) critical value is listed for the two-tail test.

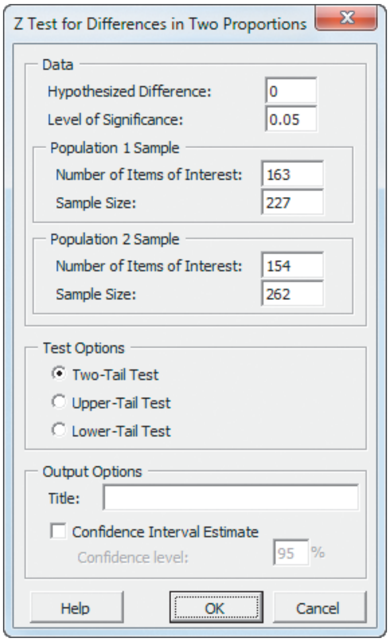
	A	B	C
1	t-Test: Paired Two Sample for Means		
2			
3		*Bookstore*	*Online*
4	Mean	139.3668	126.7005
5	Variance	3028.359	2704.292
6	Observations	19	19
7	Pearson Correlation	0.839615	
8	Hypothesized Mean Difference	0	
9	df	18	
10	t Stat	1.813248	
11	P(T<=t) one-tail	0.043252	
12	t Critical one-tail	1.734064	
13	P(T<=t) two-tail	0.086504	
14	t Critical two-tail	2.100922	

EG10.3 COMPARING the PROPORTIONS of TWO INDEPENDENT POPULATIONS

Z Test for the Difference Between Two Proportions

PHStat2 Use ***Z* Test for Differences in Two Proportions** to perform this *Z* test. For example, to perform the Figure 10.13 *Z* test for the hotel guest satisfaction survey on page 388, select **PHStat ➔ Two-Sample Tests (Summarized Data) ➔ *Z* Test for Differences in Two Proportions**. In the procedure's dialog box (shown below):

1. Enter **0** as the **Hypothesized Difference**.
2. Enter **0.05** as the **Level of Significance**.
3. For the Population 1 Sample, enter **163** as the **Number of Items of Interest** and **227** as the **Sample Size**.
4. For the Population 2 Sample, enter **154** as the **Number of Items of Interest** and **262** as the **Sample Size**.
5. Click **Two-Tail Test**.
6. Enter a **Title** and click **OK**.



In-Depth Excel Use the **COMPUTE worksheet** of the ***Z* Two Proportions workbook**, shown in Figure 10.13 on page 388, as a template for performing the two-tail *Z* test for the difference between two proportions. The worksheet contains data for the hotel guest satisfaction survey. In cells B21 and B22 respectively, the worksheet uses **NORMSINV(*level of significance*/2)** and **NORMSINV(1 - *level of significance*/2)** to compute the lower and upper critical values. In cell B23, the worksheet uses the expression **2 * (1 - NORMSDIST(*absolute value of the Z test statistic*))** to compute the *p*-value.

For other problems, change the values in cells B4, B5, B7, B8, B10, and B11 as necessary. Use the similar **COMPUTE_LOWER** or **COMPUTE_UPPER worksheets** in the same workbook as templates for performing one-tail separate-variance *t* tests.

Confidence Interval Estimate for the Difference Between Two Proportions

PHStat2 Use the *PHStat2* instructions for the *Z* test for the difference between two proportions. In step 6, also check

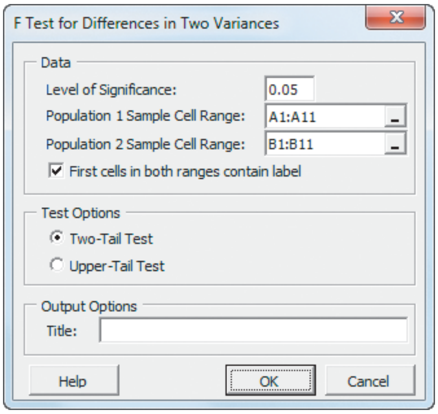
Confidence Interval Estimate and enter a **Confidence Level** in its box, in addition to entering a **Title** and clicking **OK**.

In-Depth Excel Use the *In-Depth Excel* instructions for the Z test for the difference between two proportions. The worksheets in the **Z Two Proportions workbook** include a confidence interval estimate for the difference between two means in the cell range D3:E16.

EG10.4 *F* TEST for the RATIO of TWO VARIANCES

PHStat2 Use **F Test for Differences in Two Variances** to perform this F test. For example, to perform the Figure 10.14 F test for the BLK Cola sales data on page 394, open to the **DATA worksheet** of the **COLA workbook**. Select **PHStat ➔ Two-Sample Tests (Unsummarized Data) ➔ F Test for Differences in Two Variances**. In the procedure's dialog box (shown below):

1. Enter **0.05** as the **Level of Significance**.
2. Enter **A1:A11** as the **Population 1 Sample Cell Range**.
3. Enter **B1:B11** as the **Population 2 Sample Cell Range**.
4. Check **First cells in both ranges contain label**.
5. Click **Two-Tail Test**.
6. Enter a **Title** and click **OK**.



For problems that use summarized data, select **PHStat ➔ Two-Sample Tests (Summarized Data) ➔ F Test for Differences in Two Variances**. In that procedure's dialog box, enter the level of significance and the sample size and sample variance for each sample.

In-Depth Excel Use the **COMPUTE worksheet** of the **F Two Variances workbook**, shown in Figure 10.14 on page 394, as a template for performing the two-tail F test for the ratio of two variances. The worksheet contains data and formulas for using the unsummarized data for the BLK Cola example. In cell B18, the worksheet uses **FINV(*level of significance* / 2, *population 1 sample degrees of freedom*, *population 2 sample degrees of freedom*)** to compute the upper critical value and in cell B19 uses the equivalent of the expression **2 * FDIST(*F test statistic*, *population 1 sample degrees of freedom*, *population 2 sample degrees of freedom*)** to compute the p-value.

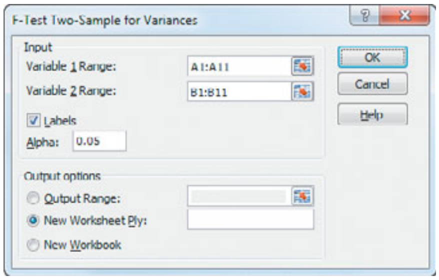
For other problems using unsummarized data, paste the unsummarized data into columns A and B of the **DATACOPY worksheet**. For summarized data, replace the COMPUTE worksheet formulas in cell ranges B6:B7 and B9:B10 with the sample statistics and ignore the DATACOPY worksheet. Use the similar **COMPUTE_UPPER worksheet** in the same workbook as a template for performing the upper-tail test.

Analysis ToolPak Use the **F-Test Two-Sample for Variances** procedure to perform the F test for the difference between two variances for unsummarized data. For example, to create results equivalent to those in the Figure 10.14 F test for the BLK Cola sales data on page 394, open to the **DATA worksheet** of the **COLA workbook** and:

1. Select **Data ➔ Data Analysis**.
2. In the Data Analysis dialog box, select **F-Test Two-Sample for Variances** from the **Analysis Tools** list and then click **OK**.

In the procedure's dialog box (shown below):

3. Enter **A1:A11** as the **Variable 1 Range** and enter **B1:B11** as the **Variable 2 Range**.
4. Check **Labels** and enter **0.05** as **Alpha**.
5. Click **New Worksheet Ply**.
6. Click **OK**.



Results (shown below) appear in a new worksheet and include only the one-tail test p-value (0.124104), which must be doubled for the two-tail test shown in Figure 10.14 on page 394.

	A	B	C
1	F-Test Two-Sample for Variances		
2			
3		*Normal*	*End-Aisle*
4	Mean	50.3	72
5	Variance	350.6778	157.3333
6	Observations	10	10
7	df	9	9
8	F	2.228884	
9	P(F<=f) one-tail	0.124104	
10	F Critical one-tail	3.178893	

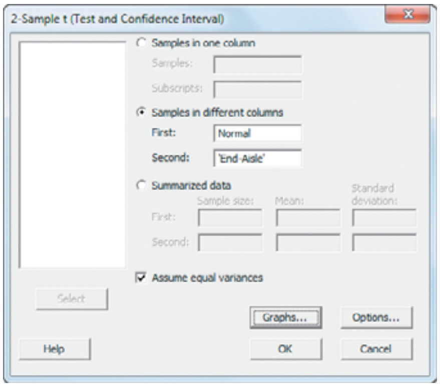
CHAPTER 10 MINITAB GUIDE

MG10.1 COMPARING the MEANS of TWO INDEPENDENT POPULATIONS

Pooled-Variance *t* Test for the Difference Between Two Means

Use **2-Sample t** to perform the pooled-variance *t* test. For example, to perform the Figure 10.3 pooled-variance *t* test for the BLK Cola data shown on page xx, open to the **Cola worksheet**. Select **Stat ➔ Basic Statistics ➔ 2-Sample t**. In the 2-Sample t (Test and Confidence Interval) dialog box (shown below):

1. Click **Samples in different columns** and press **Tab**.
2. Double-click **C1 Normal** in the variables list to add **Normal** to the **First** box.
3. Double-click **C2 End-Aisle** in the variables list to add **'End-Aisle'** to the **Second** box.
4. Check **Assume equal variances**.
5. Click **Graphs**.



In the 2-Sample t - Graphs dialog box (not shown):

6. Check **Boxplots of data** and then click **OK**.
7. Back in the 2-Sample t (Test and Confidence Interval) dialog box, click **OK**.

For stacked data, use these replacement steps 1 through 3:

1. Click **Samples in one column**.
2. Enter the name of the column that contains the measurement in the **Samples** box.
3. Enter the name of the column that contains the sample names in the **Subscripts** box.

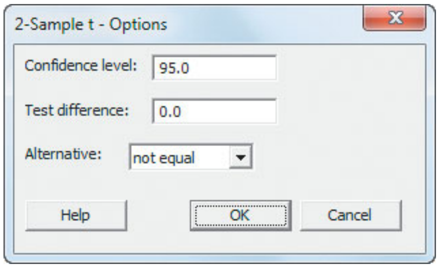
Confidence Interval Estimate for the Difference Between Two Means

Use the instructions for the pooled-variance *t* test, replacing step 7 with these steps 7 through 12:

7. Back in the 2-Sample t (Test and Confidence Interval) dialog box, click **Options**.

In the 2-Sample t - Options dialog box (shown below):

8. Enter **95.0** in the **Confidence level** box.
9. Enter **0.0** in the **Test difference** box.
10. Select **not equal** from the **Alternative** drop-down list (to perform the two-tail test).
11. Click **OK**.
12. Back in the 2-Sample t (Test and Confidence Interval) dialog box, click **OK**.



To perform a one-tail test, select **less than** or **greater than** in step 10.

t Test for the Difference Between Two Means, Assuming Unequal Variances

Use the instructions for the pooled-variance *t* test with this replacement step 4:

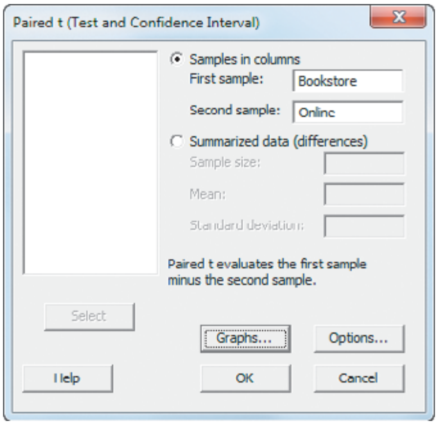
4. Clear **Assume equal variances.**

MG10.2 COMPARING the MEANS of TWO RELATED POPULATIONS

Paired *t* Test

Use **Paired t** to perform the paired *t* test. For example, to perform the Figure 10.9 paired *t* test for the textbook price data on page 381, open to the **BookPrices worksheet**. Select **Stat ➔ Basic Statistics ➔ Paired t**. In the Paired *t* (Test and Confidence Interval) dialog box (shown on the top of page 412):

1. Click **Samples in columns** and press **Tab**.
2. Double-click **C3 Bookstore** in the variables list to enter **Bookstore** in the **First sample** box.
3. Double-click **C4 Online** in the variables list to enter **Online** in the **Second sample** box.
4. Click **Graphs**.



In the Paired t - Graphs dialog box (not shown):

5. Check **Boxplots of differences** and then click **OK**.
6. Back in the Paired *t* (Test and Confidence Interval) dialog box, click **OK**.

Confidence Interval Estimate for the Mean Difference

Use the instructions for the paired *t* test, replacing step 6 with these steps 6 through 11:

6. Back in the Paired *t* (Test and Confidence Interval) dialog box, click **Options**.

In the Paired t - Options dialog box (not shown):

7. Enter **95.0** in the **Confidence level** box.
8. Enter **0.0** in the **Test difference** box.
9. Select **not equal** from the **Alternative** drop-down list (to perform the two-tail test).
10. Click **OK**.
11. Back in the Paired *t* (Test and Confidence Interval) dialog box, click **OK**.

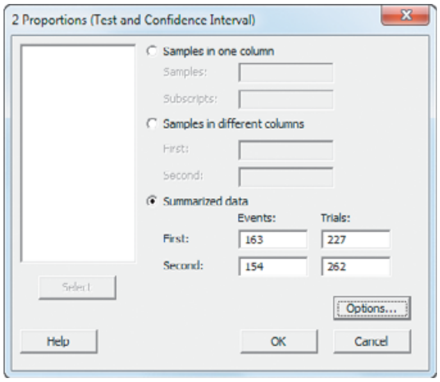
To perform a one-tail test, select **less than** or **greater than** in step 9.

MG10.3 COMPARING the PROPORTIONS of TWO INDEPENDENT POPULATIONS

Z Test for the Difference Between Two Proportions

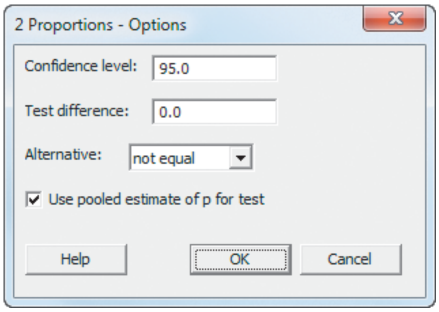
Use **2 Proportions** to perform this *Z* test. For example, to perform the Figure 10.13 *Z* test for the hotel guest satisfaction survey on page 388, select **Stat → Basic Statistics → 2 Proportions**. In the 2 Proportions (Test and Confidence Interval) dialog box (shown at the top of the right column):

1. Click **Summarized data**.
2. In the **First** row, enter **163** in the **Events** box and **227** in the **Trials** box.
3. In the **Second** row, enter **154** in the **Events** box and **262** in the **Trials** box.
4. Click **Options**.



In the 2 Proportions - Options dialog box (shown below):

5. Enter **95.0** in the **Confidence level** box.
6. Enter **0.0** in the **Test difference** box.
7. Select **not equal** from the **Alternative** drop-down list.
8. Check **Use pooled estimate of p for test**.
9. Click **OK**.
10. Back in the 2 Proportions (Test and Confidence Interval) dialog box, click **OK**.



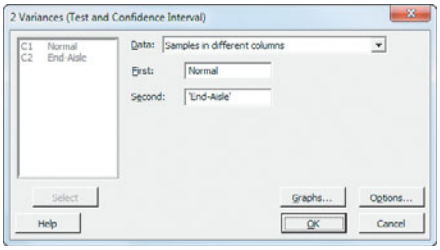
Confidence Interval Estimate for the Difference Between Two Proportions

Use the "*Z* Test for the Difference Between Two Proportions" instructions above to compute the confidence interval estimate.

MG10.4 *F* TEST for the RATIO of TWO VARIANCES

Use **2 Variances** to perform this *F* test. For example, to perform the Figure 10.14 *F* test for the BLK Cola sales data on page 394, open to the **COLA worksheet**. Select **Stat → Basic Statistics → 2 Variances**. In the 2 Variances (Test and Confidence) dialog box (shown on page 413):

1. Select **Samples in different columns** from the **Data** drop-down list and press **Tab**.
2. Double-click **C1 Normal** in the variables list to add **Normal** to the **First** box.
3. Double-click **C2 End-Aisle** in the variables list to add **'End-Aisle'** to the **Second** box.
4. Click **Graphs**.



In the 2 Variances - Graph dialog box (not shown):

5. Clear all check boxes
6. Click **OK**.
7. Back in the 2 Variances (Test and Confidence) dialog box, click **OK**.

For summarized data, select **Sample standard deviations** or **Sample variances** in step 1 and enter the sample size and the sample statistics for the two variables in lieu of steps 2 and 3. For stacked data, use these replacement steps 1 through 3:

1. Select **Samples in one column** from the **Data** drop-down list.
2. Enter the name of the column that contains the measurement in the **Samples** box.
3. Enter the name of the column that contains the sample names in the **Subscripts** box.

If you use an older version of Minitab, you will see a 2 Variances dialog box instead of the 2 Variances (Test and Confidence) dialog box shown and described in this section. This older dialog box is similar to the dialog boxes shown in the preceding three sections: You click either **Samples in different columns** or **Summarized data** and then make entries similar to the ones described in the previous sections. The results created using older versions differ slightly from the Minitab results shown in Figure 10.14.